

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 16:05:16 ; Search time 6071.42 Seconds

(without alignments)
11868.013 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atgagatcgctgagcttact.....gaatctgctcaagctcagc 1893

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2899	3 AK004737	AK004737 Mus muscu
2	1893	100.0	3098	3 AK033672	AK033672 Mus muscu
3	1857.4	98.1	2869	3 AK034082	AK034082 Mus muscu
4	1773	93.7	1896	9 AV399041	AV399041 Mus muscu
5	1355.4	71.6	3443	3 HSM801859	AV399041 Homo sapi
6	1261	66.6	1887	9 AV399039	AV399039 Homo sapi
7	948.4	50.1	1887	9 AV399040	AV399040 Pan trogl
8	925.4	48.9	7004	3 CR749209	CR749209 Homo sapi
9	683.6	36.1	726	6 CB248251	CB248251 UT-M-FD0-
10	672	35.5	672	6 CB248251	CB248251 UT-M-FD0-
11	652.4	34.5	655	7 CK781308	CK781308 UI-M-G10-
12	651.6	34.4	805	4 BG174288	BG174288 G02334589
13	646	34.1	646	4 BG694881	BG694881 NISC 1V09
14	645.8	34.1	948	2 BF789245	BF789245 G02105144
15	640	33.8	795	6 CA319312	CA319312 UT-M-FW0-
16	627.6	33.2	782	7 CN261003	CN261003 170004247
17	614.8	32.5	1069	5 BM927376	BM927376 AGENCOURT
18	605.8	32.0	1022	5 BQ062868	BQ062868 AGENCOURT
19	603	31.9	893	5 BF785290	BF785290 G02108490
20	596	31.5	705	7 CF726196	CF726196 UT-M-GZ0-
21	586.4	31.0	1005	5 BQ060729	BQ060729 AGENCOURT
22	586.4	31.0	1029	5 BQ055908	BQ055908 AGENCOURT
23	578	30.5	722	4 B1143493	B1143493 G02907673
24	574.4	30.3	1012	5 BQ064231	BQ064231 AGENCOURT

25	562.8	29.7	713	2	BB625283
26	558.8	29.5	667	4	B1648260
27	554.8	29.3	558	8	BH116351
28	552.2	29.2	1042	5	BM909401
29	547	28.9	939	2	BE912458
30	541.4	28.6	761	6	CD632904
31	541.4	28.6	851	5	BQ228581
32	535.2	28.3	824	7	CK483217
33	534	28.2	1003	5	BQ056806
34	532.8	28.1	890	6	CA487493
35	531.8	28.1	535	4	BG086714
36	530	28.0	1120	5	BM909505
37	529.2	28.0	760	4	BG753131
38	528	27.9	1038	5	BQ065633
39	526.2	27.8	1047	5	BQ064127
40	524.8	27.7	865	4	BG172224
41	522.4	27.6	616	4	B1647932
42	521	27.5	818	7	CK483904
43	508	26.8	556	6	CA535696
44	506.4	26.8	1044	5	BQ065997
45	505.8	26.7	639	6	CD632908

ALIGNMENTS

RESULT 1	AK004737	2899 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK004737				
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013B2 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (BC 2.7.1.-) (Homo sapiens), full insert sequence.				
ACCESSION	AK004737				
VERSION	AK004737.2	GI:26334437			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
REFERENCE	99279253				
REFERENCE	10349636				
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	20499374				
REFERENCE	11042159				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, T., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuno, H., Sakaue, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)				
REFERENCE	20530913				
REFERENCE	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
REFERENCE	5				

URL: <http://fantom.gsc.riken.jp/>.

FEATURES

source

Location/Qualifiers

1..3098

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM.DB:9130215K18"

/db_xref="taxon:10090"

/clone="9130215K18"

/sex="male"

/tissue_type="cecum"

/clone_id="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

93..1988

/note="unamed protein product; putative

weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE

KIA0537 (BC 2.7.1.-) [Homo sapiens] (SWISSPROT 060285,

evidence: FASTV, 56.8%ID, 98.6%length, match=1849)"

/codon_start=1

/protein_id="PAC28421.1"

/db_xref="GI:26329365"

/translation="MESVALLQRPSPASALASBPAPLADGLIKSPRLMKQAV

KRHHKTNLHRYEFLFTLGKTYKAKAESGRLVAIKIKDKIKDEODILHTR

REIEMSLNHPHIIAHEVENSKIVIMYASRGDLVDYISERPLSERDHPF

ROIYSALHYCHONGIIVHRDLKLENIILDANGNIKIADPGLSNLVHKGKPIOTPGSP

YASPEIYNGKPYGPEVDWSIGVLIVIVGTMPPGQPHKTLYKQISGAREPPE

PSDACGLIRHLIMYNTPRRATLEBVASHMWNMGYTTTGVEQELKRGHPSGREGA

SMADWKRSPLEENAKVCSFFKQHVPGGSTVPELREQSHLSKSKNDMAQNLQ

GDPAEDTSRPPKSLKPKGIILKKSSTSGEVDPEQIBPVPDGPVPVAVSLI

PKGIILKKSREGYSSPEPSGELIASDVFGADPEVQKSPQASGLLHRKGI

LKNKFSRTALEGTSTPSGSLDQLASHPAARPSRPSGASDSTLISSESPDLDL

PERLPETPLRGCVSDNLRLGLEPSPBGLKRWQESLGDSCFSLTDOQVTAAYROL

GICSKLS"

3082..3087

polyA_signal

/note="putative"

polyA_site

3098

/note="putative"

ORIGIN

Query Match

Best Local Similarity

Matches 1893; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

100.0%; Score 1893; DB 3; Length 3098;

100.0%; Pred. No. 0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

Db 513 CTGTATGATTACATCAAGTAGCGGGCCAGCGCTGAGTAGGGGACGCCAGGCAATTCTTC 572

Qy 481 CGACAGATCTGTCTGCGCTTGCATCTACTGCGCACGAGATCGTTCAACCGAGATCTC 540

Db 573 CGACAGATCTGTCTGCGCTTGCATCTACTGCGCACGAGATCGTTCAACCGAGATCTC 632

Qy 541 AAGCTGGAACATCCCTTCTAGATGCGAATGGAACATCAAGATGCTGACTTTGGGCTC 600

Db 633 AAGCTGGAACATCCCTTCTAGATGCGAATGGAACATCAAGATGCTGACTTTGGGCTC 692

Qy 601 TCCAACTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGGAGCCCTCTTACGCC 660

Db 693 TCCAACTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGGAGCCCTCTTACGCC 752

Qy 661 TCCGCTGAGATGATCAACGGGAAACCTTATGTGGGCCAGAGGTGACACAGCTGCTCTG 720

Db 753 TCCGCTGAGATGATCAACGGGAAACCTTATGTGGGCCAGAGGTGACACAGCTGCTCTG 812

Qy 721 GGGCTTCTCTGTACATCCCTGTGTGATGACCAATGCTTGTGACGGGAGAGATCATAA 780

Db 813 GGGCTTCTCTGTACATCCCTGTGTGATGACCAATGCTTGTGACGGGAGATCATAA 872

Qy 781 ACACTGTGAGGAAATCAATACGGGAGCTTACCGTGAAGCGGCCAAAGCGTCCAGTGC 840

Db 873 ACACTGTGAGGAAATCAATACGGGAGCTTACCGTGAAGCGGCCAAAGCGTCCAGTGC 932

Qy 841 TGTGGCTGATCCGGTGTGCTTTAAATGTGAACCCACCCGTTGGGCGACATGAGAT 900

Db 933 TGTGGCTGATCCGGTGTGCTTTAAATGTGAACCCACCCGTTGGGCGACATGAGAT 992

Qy 901 GTAGCAATGATGTGGGTGCACTGGGGTTTACCAACCGGATCGGGGAAACAGGAAGC 960

Db 993 GTAGCAATGATGTGGGTGCACTGGGGTTTACCAACCGGATCGGGGAAACAGGAAGC 1052

Qy 961 CTGCTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGCGCTTCAATGCGGAGCTGTTA 1020

Db 1053 CTGCTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGCGCTTCAATGCGGAGCTGTTA 1112

Qy 1113 CTGCTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGCGCTTCAATGCGGAGCTGTTA 1172

Db 1081 CAGTGTGGGAGGTGGAAGCACTGACCTGGGCTGGAAGGGAACATTTCTTAAGAG 1140

Qy 1173 CAGTGTGGGAGGTGGAAGCACTGACCTGGGCTGGAAGGGAACATTTCTTAAGAG 1232

Db 1141 TCCGAAAGGAGATGATCATGTGCTCAAAATCTCAAGGTGACCCGCTGAGATACCTC 1200

Qy 1233 TCCGAAAGGAGATGATCATGTGCTCAAAATCTCAAGGTGACCCGCTGAGATACCTC 1292

Db 1201 TCTGCGCTTGGCAAGACAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCT 1260

Qy 1293 TCTGCGCTTGGCAAGACAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCT 1352

Db 1261 ACCTGCTCAGGGAGGTGACAGAGAACCTTCAAGAACTCAAGACGGTGTCTGATCTCCA 1320

Qy 1353 ACCTGCTCAGGGAGGTGACAGAGAACCTTCAAGAACTCAAGACGGTGTCTGATCTCCA 1412

Db 1321 GGGCAGGCTGTCCCTGCTGATCCCTGTGCTCCCAAGGAAAGGCAATCTTAAAGATCTCGA 1380

Qy 1413 GGGCAGGCTGTCCCTGCTGATCCCTGTGCTCCCAAGGAAAGGCAATCTTAAAGATCTCGA 1472

Db 1381 CAGCGTGAATCTGTGTTACTACTCTCTCCAGAGCCGAGAGTCTGGGAACTTTAGAC 1440

Qy 1473 CAGCGTGAATCTGTGTTACTACTCTCTCCAGAGCCGAGAGTCTGGGAACTTTAGAC 1532

Db 1441 GCGAGTGAATGTTTGTGATGGGAGCCCGTGGAGAGAGTCTCAAGGCTTCAAGG 1500

Qy 1533 GCGAGTGAATGTTTGTGATGGGAGCCCGTGGAGAGAGTCTCAAGGCTTCAAGG 1592

Db 1501 CTGCTCTCCACCGCAAGGCAATTTCAAACTCAAGTGAAGTCTCCGCAAGGCTTCA 1560

Qy 1593 CTGCTCTCCACCGCAAGGCAATTTCAAACTCAAGTGAAGTCTCCGCAAGGCTTCA 1652

QY 1561 GAAGGCACTACCCCTAGCACCCTTGGCTCCCTGAGCAACTGGCTCTCTCCATCTGCA 1620
1563 GAAGGCACTACCCCTAGCACCCTTGGCTCCCTGAGCAACTGGCTCTCTCCATCTGCA 1712
QY 1621 GCCCGGCCGAGCCGCCCTTCAAGGCGCTGATGATGAGACAGACTCTGTCTCCGAGTCC 1680
1713 GCCCGGCCGAGCCGCCCTTCAAGGCGCTGATGATGAGACAGACTCTGTCTCCGAGTCC 1772
QY 1681 TTGACCAATTTGACTTGGCTGGAACGCTTCTCCGAAACCCCACTGAGGGGCTGATGCT 1740
1773 TTGACCAATTTGACTTGGCTGGAACGCTTCTCCGAAACCCCACTGAGGGGCTGATGCT 1832
QY 1741 GTGACCACTGAGGGGCTTGAAGCAGCTCTCCAGAGGCTGAAAGCTGAGTGGCAG 1800
1833 GTGACCACTGAGGGGCTTGAAGCAGCTCTCCAGAGGCTGAAAGCTGAGTGGCAG 1892
QY 1801 GAATCTTTGGGGGATGAGCTGCTTTCTCTGACAGACTGCGCAAGGAGTGAAGCCTTAC 1860
1893 GAATCTTTGGGGGATGAGCTGCTTTCTCTGACAGACTGCGCAAGGAGTGAAGCCTTAC 1952
QY 1861 AGACAGCCCTAGAGATCTGCTCAAGCTCAGC 1893
1953 AGACAGCCCTAGAGATCTGCTCAAGCTCAGC 1985
Db
RESULT 3
LOCUS AK034082 2869 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
cDNA library, clone:9330154N24 product:weakly similar to
PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)
[Homo sapiens], full insert sequence.
AK034082
VERSION AK034082.1 GI:26329672
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
METHOD Enzymol. 303, 19-44 (1999)
PUBMED 99279253
10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kizunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequence
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2869)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imclanti, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
Location/Qualifiers
1. 2869
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9330154N24"
/db_xref="taxon:10090"
/clone="9330154N24"
/sex="male"
/tissue type="dienecephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
98.2017
/note="unnamed protein product; putative
weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|060285,
evidence: FASTV, 56.8%id, 98.6%length, match=1849)"
/codon_start=1
/protein_id="BAC28575.1"
/db_xref="GI:26329673"
/translation="MESGALLQRPSPQAPSALASBSARPLADGLIKSPYLMKQAV
KRHHKHNLRHREPELTGKGYGKAKRESGRVIAIKSKIDKIQEODLHNR
REIENGSLNHPHIIAHEVGRSLVIVFNSGRILVIMVYAGRDLDVYSISRRPS
ERARHFRFIVGSAHLYCHONGYVHRDILKENTILIDANKIKADRELSNLYKRGTL
QTGCSFLVSPFIVNKPVYGEVDSWISGVLLYIIVHGTMPDGDHDLTVQJIN
GASDFKPSADAGLIRMLIMVPTRAVLEDAVSHWVMVGTYTGGDEBALREGH
PSDFPGASADAGLIRRSRSLILNKGAVCSFFQHYVGGSTVGLRGLSKSRK
NDMAONLQGPADTSSRPGKSKILPKGLIKKSSYSGEVOBPOELRHPVPTPOG
PVAVALILPKRGIKTSRQESGIVYSPPRSEGLDADSVPSGDPVBRKPSQASG
ILHRRKIIILNKGFRSTALREGTTPSRGSLDQASHPARAPRSPGAVSESDILS
ESFDQDPERLPETPLRGCVSDNLKGLQRPSEGLKRWQBSLSDSCFSLTDCQEV
TAAYRQALGICSKLS"

ORIGIN

Query Match 98.1%; Score 1857.4; DB 3; Length 2869;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 ATGAGTGGTGGCTTACTTCAGCGCCGAGCAGGCTCCCTCGGCTCGCCTGACC 60

Db 98 ATGAGTTCGGGGGCTTACTTCACGCGCCGACGAGGCTCCCTGGGCTCCGCGCTTGGCC 157
Qy 61 TCGAGAGCCGCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAAACCTTGATGAAG 120
Db 158 TCGAGAGCCGCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAAACCTTGATGAAG 217
Qy 121 AAGAGGCGGGAAGCGGCGACCATCAAAACAACCTGGGCGACCGCTAAGATTTCTG 180
Db 218 AAGAGGCGGGAAGCGGCGACCATCAAAACAACCTGGGCGACCGCTAAGATTTCTG 277
Qy 181 GAGAGCTGGGCAAGGCGACCTACGAGAGGTGAAGAGGACGAGAGAGCTCGGGCGT 240
Db 278 GAGAGCTGGGCAAGGCGACCTACGAGAGGTGAAGAGGACGAGAGAGCTCGGGCGT 337
Qy 241 CTGGTGGCCATCAAGTTCATCAGAAAGACAAATCAAGATGACAGAGATCTGCTGAC 300
Db 338 CTGGTGGCCATCAAGTTCATCAGAAAGACAAATCAAGATGACAGAGATCTGCTGAC 397
Qy 301 ATTACGAGGAGATTTGATGATGCTTCACTCAACCCCGCATCATTTGCCATGCAT 360
Db 398 ATTACGAGGAGATTTGATGATGCTTCACTCAACCCCGCATCATTTGCCATGCAT 457
Qy 361 GAA-----GTGTTGAGATAGCAGCAGATTTGATTTGTC 396
Db 458 GAAATGGGAGATCAGCTTATGATGATGCTTTGAGATAGCAGCAGATTTGATTTGTC 517
Qy 397 ATGAGATATGCGACGCGAGCGATCTGATGATTTACATCAGTGAACGCGCATGAT 456
Db 518 ATGAGATATGCGACGCGAGCGATCTGATGATTTACATCAGTGAACGCGCATGAT 577
Qy 457 GAGGAGGACCGCGAGGATTTCTTCCGACAGATGCTGCTGCGCTGACATGCGACACG 516
Db 578 GAGGAGGACCGCGAGGATTTCTTCCGACAGATGCTGCTGCGCTGACATGCGACACG 637
Qy 517 AACGGATTCGTTACCGAGATCTCAAGTGAAGAAACCTCTTCAAGTGGCAATGGAAC 576
Db 638 AACGGATTCGTTACCGAGATCTCAAGTGAAGAAACCTCTTCAAGTGGCAATGGAAC 697
Qy 577 ATCAAGATTCGTTACCTTGGCTCTGCAACCTGTACCAAAAGGCAAGTTCTTCAGACG 636
Db 698 ATCAAGATTCGTTACCTTGGCTCTGCAACCTGTACCAAAAGGCAAGTTCTTCAGACG 757
Qy 637 TTCTGTGGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACGCGAAACCTTATGTGGC 696
Db 758 TTCTGTGGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACGCGAAACCTTATGTGGC 817
Qy 697 CCAGAGGTGACAGCTGATCTGAGGCTTCTCTGTACATCTGATGACATGGACCATG 756
Db 818 CCAGAGGTGACAGCTGATCTGAGGCTTCTCTGTACATCTGATGACATGGACCATG 877
Qy 757 CCTTTGACGGGCGAGATCAAAACACTGTGAAGAAATCAGTAAACGCGGCTTACCGT 816
Db 878 CCTTTGACGGGCGAGATCAAAACACTGTGAAGAAATCAGTAAACGCGGCTTACCGT 937
Qy 817 GAGCGCGCCAAAGCGCTCCGATGCTGTGGCTGATCCGCTGCTGTTAATGTGAACCC 876
Db 938 GAGCGCGCCAAAGCGCTCCGATGCTGTGGCTGATCCGCTGCTGTTAATGTGAACCC 997
Qy 877 ACCGATGGGCGACACTGAGAGATGTAAGCAGTATGTGGGCTCACTGGGGTTAACCC 936
Db 998 ACCGATGGGCGACACTGAGAGATGTAAGCAGTATGTGGGCTCACTGGGGTTAACCC 1057
Qy 937 ACCGATGGGCGACACTGAGAGCCTCGTGAAGGCTGGGACCTTATGTGATTTGGC 996
Db 1058 ACCGATGGGCGACACTGAGAGCCTCGTGAAGGCTGGGACCTTATGTGATTTGGC 1117
Qy 997 CGGCGCTCATGCGGAGCTGTTAGTGTGCTCTCGCGCCCTCTGAGAGATGAGCC 1056
Db 1118 CGGCGCTCATGCGGAGCTGTTAGTGTGCTCTCGCGCCCTCTGAGAGATGAGCC 1177
Qy 1057 AAGGTGCGAGCTTCTTCAAGCAGCAGTGCAGGAGGTGAAGCAGTATACCTGGGCTG 1116

Db 1178 AAGGTGCGAGCTTCTTCAAGCAGCAGTGCAGGAGGTGAAGCAGTATACCTGGGCTG 1237
Qy 1117 GAGCGCGACATTTCTTAAAGAGTCCCGAAAGAGATGATGCTCAAAATCTGCA 1176
Db 1238 GAGCGCGACATTTCTTAAAGAGTCCCGAAAGAGATGATGCTCAAAATCTGCA 1297
Qy 1177 GGTGACCGGCTGAGATACCTTCTGCGCTGGGCAAGGCGCTTAACTTCGAA 1236
Db 1298 GGTGACCGGCTGAGATACCTTCTGCGCTGGGCAAGGCGCTTAACTTCGAA 1357
Qy 1237 GGCATTTCAAGAAAAAGTCTCTACCTCGTACGAGGAGGTACAGAGACCTTCAGAA 1296
Db 1358 GGCATTTCAAGAAAAAGTCTCTACCTCGTACGAGGAGGTACAGAGACCTTCAGAA 1417
Qy 1297 CTGACCGGCTGATTAATCTCAGGAGCAGCTGCTGCTGATATCTGCTCCAGG 1356
Db 1418 CTGACCGGCTGATTAATCTCAGGAGCAGCTGCTGCTGATATCTGCTCCAGG 1477
Qy 1357 AAAGGATCTCTTAAAGATCTGACAGCGGAAATCTGTTACTACCTCTCCAGAGCC 1416
Db 1478 AAAGGATCTCTTAAAGATCTGACAGCGGAAATCTGTTACTACCTCTCCAGAGCC 1537
Qy 1417 AGCGAGTCTGGGAACTCTTACAGCGCAGATGATGTTGTGAGTGGGAGCCCGTGAAG 1476
Db 1538 AGCGAGTCTGGGAACTCTTACAGCGCAGATGATGTTGTGAGTGGGAGCCCGTGAAG 1597
Qy 1477 CAGAGTCTCAGAGGCTTCAAGGCTCTCTCAGCGACCGAGGCAATCTCAATCAAT 1536
Db 1598 CAGAGTCTCAGAGGCTTCAAGGCTCTCTCAGCGACCGAGGCAATCTCAATCAAT 1657
Qy 1537 GCGAAGTCTCCGCGACAGCTTGAAGGACATACCTTGAAGCCTTGGCTCCCTGAGC 1596
Db 1658 GCGAAGTCTCCGCGACAGCTTGAAGGACATACCTTGAAGCCTTGGCTCCCTGAGC 1717
Qy 1597 CAAGTGGCTCTCCATCTCCATCTGAGCGCCGCGCGCGCGCGCGCGCGCGCGCGCG 1656
Db 1718 CAAGTGGCTCTCCATCTCCATCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1777
Qy 1657 GACAGATCTGCTCTCCAGATCTTGAACCAATGAGATCTGCTGAACGCTTCCGAA 1716
Db 1778 GACAGATCTGCTCTCCAGATCTTGAACCAATGAGATCTGCTGAACGCTTCCGAA 1837
Qy 1717 ACCCACTGAGGGGCTGCTGCTGCTGAGCAACCTGAGGGGCTTGAAGCGCTCCCTCA 1776
Db 1838 ACCCACTGAGGGGCTGCTGCTGCTGAGCAACCTGAGGGGCTTGAAGCGCTCCCTCA 1897
Qy 1777 GAGGCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1836
Db 1898 GAGGCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1957
Qy 1837 TGCAGAGAGGTGATCAGCTTACAGACAAAGCCTTGAATGCTCAAGCTCAGC 1893
Db 1958 TGCAGAGAGGTGATCAGCTTACAGACAAAGCCTTGAATGCTCAAGCTCAGC 2014

RESULT 4
AY399041 1896 bp DNA linear GSS 12-DEC-2003
LOCUS AY399041
DEFINITION Mus musculus HCM0091 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY399041
VERSION AY399041.1 GI:39755030
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1896)
REFERENCE
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaratnal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trilos
Science 302 (5652), 1960-1963 (2003)
14671302
PUBMED
2 (bases 1 to 1896)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejalrajal, A.,
Todd, M.A., Tanenbaum, D.M., Clevello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sainsbury, D.J.,
Adams, M.D. and Cargill, M.
TITLE
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 1896
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1896
/locus_tag="HCOM091"
ORIGIN
Query Match 93.7%; Score 1773; DB 9; Length 1896;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1773; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 1 ATGAGTGGTGGCTTACTTCAGAGCCCGAGAGGAGGCTCCCTGGGCTCGGCTGGCC 60
DB 1 ATGAGTGGTGGCTTACTTCAGAGCCCGAGAGGAGGCTCCCTGGGCTCGGCTGGCC 60
QY 61 TCGAGAGCGCCCGCGCGCTGGCGAGCGGGCTCATCAAGTCGCTAAACCTGTGTAAG 120
DB 61 TCGAGAGCGCCCGCGCGCTGGCGAGCGGGCTCATCAAGTCGCTAAACCTGTGTAAG 120
QY 121 AACGAGCGGTGAAAGCGGCAATCAACAACAACCTGGCGCAACGCTCAAGTTCCTG 180
DB 121 AACGAGCGGTGAAAGCGGCAATCAACAACAACCTGGCGCAACGCTCAAGTTCCTG 180
QY 121 AAGCAGCGGTGAAAGCGGCAATCAACAACAACCTGGCGCAACGCTCAAGTTCCTG 180
DB 121 AAGCAGCGGTGAAAGCGGCAATCAACAACAACCTGGCGCAACGCTCAAGTTCCTG 180
QY 181 GAGAGCGTGGGCAAGGAGCCTTACGGGAGGTGAAAGGAGCAGAGAGCTGGGGCGT 240
DB 181 GAGAGCGTGGGCAAGGAGCCTTACGGGAGGTGAAAGGAGCAGAGAGCTGGGGCGT 240
QY 241 CTGGTGGCCATCAAGTTCATCAGAAAGACAAATCAAGATGAGAGGATTCCTGAC 300
DB 241 CTGGTGGCCATCAAGTTCATCAGAAAGACAAATCAAGATGAGAGGATTCCTGAC 300
QY 301 ATACGAGGAGATTTGATGATGATCTTCACTCAACACCCCAATTCATTCATCAT 360
DB 301 ATACGAGGAGATTTGATGATGATCTTCACTCAACACCCCAATTCATTCATCAT 360
QY 301 ATACGAGGAGATTTGATGATGATCTTCACTCAACACCCCAATTCATTCATCAT 360
DB 301 ATACGAGGAGATTTGATGATGATCTTCACTCAACACCCCAATTCATTCATCAT 360
QY 361 GAAGTGTGGAATGACAGCAAGATTTGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GAAGTGTGGAATGACAGCAAGATTTGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CTGTATGATTAATCATGATGAGCGGCAACGCTGAGTGGAGCGGCAAGGATTTCTT 480
DB 421 CTGTATGATTAATCATGATGAGCGGCAACGCTGAGTGGAGCGGCAAGGATTTCTT 480
QY 481 CGAAGATCTGTGTGCTGCTGCACTACTGCGCAACGAGATGTTCAACGAGATCTC 540
DB 481 CGAAGATCTGTGTGCTGCTGCACTACTGCGCAACGAGATGTTCAACGAGATCTC 540
QY 541 AACCTGGAAGAACTCTTCTAGATGCAATGGAAGAACTCAAGATGCTGATTTGGGCTC 600
DB 541 AACCTGGAAGAACTCTTCTAGATGCAATGGAAGAACTCAAGATGCTGATTTGGGCTC 600
QY 541 AACCTGGAAGAACTCTTCTAGATGCAATGGAAGAACTCAAGATGCTGATTTGGGCTC 600
DB 541 AACCTGGAAGAACTCTTCTAGATGCAATGGAAGAACTCAAGATGCTGATTTGGGCTC 600
QY 601 TCCAACTGTACCAAAAGGCAAGTTCCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 660
DB 601 TCCAACTGTACCAAAAGGCAAGTTCCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 660
QY 661 TCGCTGAGATGATGACAGGAGACCTTATGTGGGCGCAGAGTGGACAGTGTCTCTG 720
DB 661 TCGCTGAGATGATGACAGGAGACCTTATGTGGGCGCAGAGTGGACAGTGTCTCTG 720
QY 661 TCGCTGAGATGATGACAGGAGACCTTATGTGGGCGCAGAGTGGACAGTGTCTCTG 720
DB 661 TCGCTGAGATGATGACAGGAGACCTTATGTGGGCGCAGAGTGGACAGTGTCTCTG 720

QY 721 GAGGTTCTCTGTACATCTGTGTGATGAGCAATGACCATGCTTTGACGGGCAAGATCATAA 780
DB 721 GAGGTTCTCTGTACATCTGTGTGATGAGCAATGACCATGCTTTGACGGGCAAGATCATAA 780
QY 781 ACACCTGTGAAGAAATCAATGATACGGGGCTTACCGTGAAGCGGCGCAAGCGCTCGATGCC 840
DB 781 ACACCTGTGAAGAAATCAATGATACGGGGCTTACCGTGAAGCGGCGCAAGCGCTCGATGCC 840
QY 841 TGTGACCTGATCCGGTGTGCTGTTAATGTGTAACCCCAACCGTGGGCGCACTGAGAGAT 900
DB 841 TGTGACCTGATCCGGTGTGCTGTTAATGTGTAACCCCAACCGTGGGCGCACTGAGAGAT 900
QY 901 GTAGCCAGTCAATTTGTGGTCAACTGGGTTACCAACCGAGTCCGGGAAACAGAAAGCC 960
DB 901 GTAGCCAGTCAATTTGTGGTCAACTGGGTTACCAACCGAGTCCGGGAAACAGAAAGCC 960
QY 961 CTGCGTGAAGGTGGGCAACCTTATGATGATCTTTGGGCGGCTTCATGCGGAGTGTGTA 1020
DB 961 CTGCGTGAAGGTGGGCAACCTTATGATGATCTTTGGGCGGCTTCATGCGGAGTGTGTA 1020
QY 1021 CGTGGCTCTCGGCGCCCTCTGTGGAATGAGCAAGGTGTGCAAGCTTCTCAAGCAG 1080
DB 1021 CGTGGCTCTCGGCGCCCTCTGTGGAATGAGCAAGGTGTGCAAGCTTCTCTCAAGCAG 1080
QY 1081 CAGTGCCTGGAGGTGGAAGCACTGTACCTGGCTGGAGCGGCAACATTTCTTAAAGAG 1140
DB 1081 CAGTGCCTGGAGGTGGAAGCACTGTACCTGGCTGGAGCGGCAACATTTCTTAAAGAG 1140
QY 1141 TCCGAAAGGAATGATCATGCTCAAAATCTCAAGATGATCCCGCTGAGATACCTCT 1200
DB 1141 TCCGAAAGGAATGATCATGCTCAAAATCTCAAGATGATCCCGCTGAGATACCTCT 1200
QY 1201 TCTGGCTCTGGCAAGAGCGCTTAAGCTTCGAAAGGCAATTCGAAAGAAAGTCTCT 1260
DB 1201 TCTGGCTCTGGCAAGAGCGCTTAAGCTTCGAAAGGCAATTCGAAAGAAAGTCTCT 1260
QY 1261 ACCTGCTCAGGGAGGTATCAGAGAGACCTTCAAGAACTCAGACCGGTCTGATATCCA 1320
DB 1261 ACCTGCTCAGGGAGGTATCAGAGAGACCTTCAAGAACTCAGACCGGTCTGATATCCA 1320
QY 1321 GGGCAAGCTGTCTCTGTGATCTCTGCTCCAAAGAGGATCTCTTAAGAGTCTGCA 1380
DB 1321 GGGCAAGCTGTCTCTGTGATCTCTGCTCCAAAGAGGATCTCTTAAGAGTCTGCA 1380
QY 1381 CAGCGTGAATCTGTGATCTCTCTCCAGAGCGAGCTGGGGAACCTTATGAC 1440
DB 1381 CAGCGTGAATCTGTGATCTCTCTCCAGAGCGAGCTGGGGAACCTTATGAC 1440
QY 1441 GCCAGTATGTTGTTGATGAGGAGACCCGCTGAGAGCAAGTCTCCACAGGCTTCAGGG 1500
DB 1441 GCCAGTATGTTGTTGATGAGGAGACCCGCTGAGAGCAAGTCTCCACAGGCTTCAGGG 1500
QY 1501 CTCCTCTCCACCGCAAGGATCTCAAACTCAATGCAAGTTCCTCCGACAGCTTAA 1560
DB 1501 CTCCTCTCCACCGCAAGGATCTCAAACTCAATGCAAGTTCCTCCGACAGCTTAA 1560
QY 1561 GAAAGCACTACCCCTGACACTTTGGCTCCCTGACCAATGAGCTCTCCATCTGCA 1620
DB 1561 GAAAGCACTACCCCTGACACTTTGGCTCCCTGACCAATGAGCTCTCCATCTGCA 1620
QY 1621 GCCGGGCGCAGCGCCCTCAGGGGCTGTGAGTGAAGAGCAGATCTCTCCAGATCC 1680
DB 1621 GCCGGGCGCAGCGCCCTCAGGGGCTGTGAGTGAAGAGCAGATCTCTCCAGATCC 1680
QY 1681 TTTGACCAATTTGATCTGCTGAAAGTCTTCCGAAACCCCACTGAGAGGCTGTGTCT 1740
DB 1681 TTTGACCAATTTGATCTGCTGAAAGTCTTCCGAAACCCCACTGAGAGGCTGTGTCT 1740
QY 1741 GTGGAACAATGAGGGGCTTGAAGAGCTCCCTCAGAGAGTGTGAAGGATGTGTGAG 1800
DB 1741 GTGGAACAATGAGGGGCTTGAAGAGCTCCCTCAGAGAGTGTGAAGGATGTGTGAG 1800

QY 1801 GAATCCTTGGGGGATAGCTGCTTTTCTTGACAAGCTCCAAAGAGTGAATCGACGCTAC 1860
Db 1801 GAATCCTTGGGGGATAGCTGCTTTTCTTGACAAGCTCCAAAGAGTGAATCGACGCTAC 1860
QY 1861 AGACAAGCCCTAGGAATCGCTCAAAAGCTCAGC 1893
Db 1861 AGACAAGCCCTAGGAATCGCTCAAAAGCTCAGC 1893

[illegible]

ORIGIN

Query Match	Best Local Match	71.6% 84.0% Matches 1562;	Score 1355.4	DB 3;	Length 3443;	
Similarity	Conservative	0;	Mismatches 276;	Indels 21;	Gaps 2	
QY	47	CCTCCGCGCTGGCTCCGGAGAGCCGCCGCCGCTGGCGGACGGGCTCAATCAAGTCGCTTA	106			
Db	162	CCACTCCCTCGGCGCGAGAGCTAACCCCGGCCGCTGGCGGAGGGCTGATCAAGTCGCCCA	221			
QY	107	AACTCTGATGAAGAAGAGCGGCTGMAAGCGGCAACATCAACAACAACCTGGGGCAAC	166			
Db	222	AGCCCTTAATAMAGAGAGCGGTGAAGCGGCAACACAAACACAACTGGCGCACCC	281			
QY	167	GCTACGAGTTCTCTGAGAGCCTGGGCAAGGCACTTACGGGAAAGTGAAGAGCAGAG	226			
Db	282	GCTACGAGTTCTCTGAGAGCCTGGGCAAGGCACTTACGGGAAAGTGAAGAGAGCGCGG	341			
QY	227	AGAGCTGGGGCGTCTGTGGCCATCAATCTCATAGAAAGACAAAATCAAAGTAGAC	286			
Db	342	AGAGCTGGGGCGCTGTGGCCATCAATCTCATAGAAAGACAAAATCAAAGTAGAC	401			
QY	287	AGATCTGCTCATATACGAGGAGATGATGATCATGTCTTCACTCAACGCCGCCACA	346			
Db	402	AAGATCTGATGCAATACGAGGAGATGATGATCATGTCTCACTCAACGCCCTCACA	461			
QY	347	TCATTTGCCATCATGAAGTGTGTTAGAAATAGACGACGAAATTTGTATTTGATGAGTATG	406			
Db	462	TCATTTGCCATCATGAAGTGTGTTAGAAACAGCACGAAAGATGTATGCTCATGAGATATG	521			
QY	407	CCAGCCGAGGGGATCTGTATGATTAATCAATCAAGTAGCGGGCCACGGCTGATGATGAGGGGACG	466			
Db	522	CCAGCCGAGGGGATCTGTATGATTAATCAATCAAGTAGCGGGCCACGGCTGATGATGAGGGGACG	581			
QY	467	CCAGGCAATTTCTTCCGACAGATCGTGTCTGCTCCGCACTACTGCGCACGAGACGGGATCG	526			
Db	582	CTAAGCAATTTCTTCCGAGAGATCGTGTCTGCTCCGCACTATTTGCAATCAGAAACAGATTG	641			
QY	527	TTCAACCGAGATCTCAAGCTGGAAAAATCCTTTCTAGATGCCAATGGAACATCAAGATTG	586			
Db	642	TCACACGAGATCTCAAGCTGGAAAAATCCTTTCTAGATGCCAATGGAATATCAAGATTG	701			
QY	587	CTGACTTTGGCTCTCCCAACCTGTACACAACAAAGGCAAGTTCTCTCAGACGTTCTGTGGGA	646			
Db	702	CTGACTTTGGCTCTCTCCCAACCTGTACACAATCAAGGCAAGTTCTCTCAGACATTTCTGTGGGA	761			
QY	647	GCCCTCTCTACGCTCGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCCCCAGAGGTGG	706			
Db	762	GCCCTCTCTACGCTCGCTGAGATGATGCAATGGGAAAGCCCTTACACAGGCCCAAGGTGG	821			
QY	707	ACACCTGTCTCTGAGGGGTTCTCCTGTACATCCGTGTGACATGAGCAATGSCCTTTAGCG	766			
Db	822	ACACCTGTCTCTGAGGGGTTCTCCTGTACATCTGTGTGACATGAGCAATGSCCTTTGATG	881			
QY	767	GGCAGGATCATAAACAACCTGGTGAAGCAAAATCACTAAACGGGGCTTAAACGTGAGCGGCCCA	826			
Db	882	GGCAGGATCATAAACAACCTGGTGAAGCAAAATCACTAAACGGGGCTTAAACGTGAGCGGCCCA	941			
QY	827	AGCGCTCGATGCTCTGTGGCTTATTCGGGTGCTGTTAATGTGTGAACCCCAACCCCTGGGG	886			
Db	942	AAACCTCTGAAGCTGTGGCTGTATTCGGGTGCTGTTAATGTGTGAACCCCAACCCCGCGGG	1001			
QY	887	CCACACTGAGAGATGTATAGCAGTCAATTGTGTGGGTCAACTGGGGGTTTACACACCGGAGTGC	946			
Db	1002	CCACACTGAGAGATGTATAGCAGTCACTGTGTGGGTCAACTGGGGGTTTACGCCACCCGAGTGG	1061			
QY	947	GGGAAACAGGAAGCCCTGCTGTAGGGGTGGGCAACCTTAATGTGTGACTTTTGGCGGGCTTCCA	1006			
Db	1062	GAGAGCAGAGAGGCTCCGCAATGAGGGGTGGGCAACCTTGCAGATGACTTGTGCCGCGCTTCCA	1121			
QY	1007	TGGGAGGACTGTTAATGTCCTCTCTCCGCGCCCTCTGAGAAATGAGAGCCAAAGTGTGCA	1066			
Db	1122	TGGGAGGACTGCTCCGGGTTCCTCTCCGCGCCCTCTCTGAGAAATGAGGGGCCAAAGTGTGCA	1181			

OY	1067	GCTTCTTCAACAGACAGCTGCGCGGAGGGTGGAAAGCACTGTACCTGTGGGTGGAGCGGCAC	1126
Db	1182	GCTTCTTCAAGCAGCATCATCTGTGGTGGGAGACACACCCCTTGCTGGAGCGGCACG	12411
OY	1127	ATTCTCTTAGAAGTCCCGAAAGAAATGACATGGCTCAAAATCTGCAGAGTACCCGG	1186
Db	1242	ATTGGCTCAAGAAATGCCGCAAGAGAAATGACATGGCCACGTCTTCCAGATGACAGG	1301
OY	1187	CTGAGGATACCTCTTCTGCGCTGGCAGAGACAGCTTAAGCTTCCGAAAGCATTC	1246
Db	1302	CTGATGACATGCCCCATGCGCTGGCAAGACATCTCAAGCTGCAAAAGGGCATTC	1361
OY	1247	AGAAAATCCCTTACCTGTCAGGGGAGGTACAGAGAGAACCTTCAGGAATCAGACCGG	1306
Db	1362	AGAGAAAGGTGTCAAGCTCTGTGAGAGAGGGTACAGAGAGACCTTCGGAGCTCAGCCCA	14211
OY	1307	TGCTGTAATCTCAGGGGAGCGCTGTCCTGTGTAATCCCTGCTCCCAAGAAAGGCATCC	1366
Db	1422	TCCCTGCGAGCCCGAGGGGAGGCTGCGC-----CGCTGCTCCCAAGAGGGCATTC	14721
OY	1367	TTAAGAAATCTCGACAGCGTGAATCTGATCTACTCTCTTCAGAGCCAGCGAGTCTG	1428
Db	1473	TCAAGAAACCCCGACAGCGCGAGTGTGACTACTACTCTCTCCGAGCCAGTGAATCTG	15321
OY	1427	GGGAATCTTGAAGCGCCAGATGTGTGTGAGATGGGAGACCCCGTGAAGCAAGTCTC	1486
Db	1533	GGGAGCTCTTGAAGCGCAGGCGACGTGTTGTGAATGGGGAATCCAGAGAGCAAGACCTC	15921
OY	1487	CACAGGCTTCAAGGGCTCTCTCCACCGCAAGGGGCACTTCTCAAACTCAATGGCAAGTCT	1546
Db	1593	CGCAAGCTTCAAGGGCTGCTCTCTCACTCGCAAAAGGCATCTCAAACTCAATGGCANATTCT	16521
OY	1547	CCCGCACAGCCTTAGAAGGCACTAACCCCTAGACCTTTGAGCTCCCTGAGCCAACTGGCT	1606
Db	1653	CCCAAGACGCTTGAAGCTCGGGGCCCCCAACACTTGGGCTCCCTGAGATGAATCGGCC	17121
OY	1607	CCTCCCATCTGCAAGCCCGGCCAGCGCCCTCAGGGGCTGTGATGAGACAGCATCC	1666
Db	1713	CACCTCGCCCCCTGGCCCGGGCCAGCCCACTCAGGGGCTGTGAGCAGAGCAGCATCC	17721
OY	1667	TGTCCTCCGAGCTCTTGAACCAATTGACCTTGCCCTGAACGTCCTCCGGAACCCCACTGA	1726
Db	1773	TGTCTCTGAGTCTTTTACAGACTGAGACTTGCTGAACGGCTCTCCAGAGCCCCCACTGCG	18321
OY	1727	GGGGCTGTGTCTGTGTGACAACTGAGGGGGCTTGAACAGCTTCCCTCAGANAG-----	1780
Db	1833	GGGGCTGTGTCTGTGTGACAACTCAAGGGGGCTTGAAGAGCCCCCTCAGAGGGCCCTG	18921
OY	1781	-----GTCTGAAGCATGTGTGGCAGGAATCTTGGGGGATAGCTGCTTTTCTGTACAG	1834
Db	1893	GAAAGCTGCTTAGAGGCGCTGCGGGCAGAGATCTTTTGGGGAGACAGCTGCTTTCCTGACAG	19521
OY	1835	ACTGCGAAGAGGTGACAGCAGACTTACCGACAGGCACTAAGAGGTGTGCTCAAGCTCAC	2011
Db	1953	ACTGCGAAGAGGTGACAGCAGACTTACCGACAGGCACTAAGAGGTGTGCTCAAGCTCAC	2011
RESULT 6			
AY399039			
LOCUS DEFINITION			
AY399039 Homo sapiens HCM0091 gene, VIRUTAL TRANSCRIPT, partial sequence, genomic survey sequence.			
ACCESSION			
AY399039			
VERSION			
AY399039.1			
KEYWORDS			
GSS.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE			
AUTHORS			
Clark, A.G., Gnanoweki, S., Nielson, R., Thomas, P., Kejarival, A.,			
Todd, M.A., Tanenbaum, D.M., Ciyello, D.R., Lu, F., Murphy, B.,			
Fertileira, S., Wang, G., Zheng, X.H., White, T.D., Smitsky, J.C.,			

TITLE	Adams,M.D. and Cargill,M.
JOURNAL	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
PUBMED	Science 302 (5652), 1960-1963 (2003)
REFERENCE	14671302
AUTHORS	2 (bases 1 to 1887) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Cavellio,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.D., Smitsky,J.O., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..1887 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1887 /locus_tag="HCM0091"
gene	
ORIGIN	
Query Match	66.6%; Score 1261; DB 9; Length 1887;
Best Local Similarity	78.4%; Pred.No.0;
Matches 1458; Conservative	0; Mismatches 380; Indels 21; Gaps 2
OY	CCTCCGCGCCTTGAGGCTCGGAGAGCGGCCGGCTGACGAGCGGGCTCAATCAAGTCGCTTA 106
Db	35 CCATCTCCTCGGCGCAGAGCTAGCCGCGCGGTGGCGAAGGGCTGATCAAGTGCCTCA 94
OY	107 AACCTCTGATGAAGAAGCAGCGCGGTGAAGCGGCACATCACAAACAACCTCGCGCAC 166
Db	95 AGCCCTTAATAAAGAGCAGCGCGGTGAAGCGGCACACACAGCACAACCTCGCGCAC 154
OY	167 GCTACGAGTTCTTGAGACGCTGGSCAAGGGCATCTACGGGAAGGTGAAGAGGCACGAG 226
Db	155 GCTACGAGTTCTTGAGACGCTGGSCAAGGGCATCTACGGGAAGGTGAAGAGGGCGGG 214
OY	227 AGAGTCGGGGCGTCTGGTGGCATCAAGTCATCAGGAAGACAAAATCAAAGATGAGC 286
Db	215 AGAGTCGGGGCGCTGTGGCATCAAGTCATCAGGAAGACAAAATCAAAGATGAGC 274
OY	287 AGAGTCGTCTCAATACGAGGAGGATTTGATCATGTCTTCACTCAACCAACCCCA 346
Db	275 AAGATCTGATGCAATACGAGGAGGATTTGATCATGTCTCACTCAACCAACCCCA 334
OY	347 TCATTGCCATCCATGAAGTGTTTGAAGATGACGACAGATGTTGTCTCATGAGATG 406
Db	335 TCATTGCCATCCATGAAGTGTTTGAAGATGACGACAGATGTTGTCTCATGAGATG 394
OY	407 CCAGCCGAGGGAGTCTGATATTAATCATCATGAGCGGCCAACGGTGAAGTAGCGGACG 466
Db	395 CCAGCCGAGGGAGCTTTATATCATCATCAACGAGCGGCAACGTCAGTAGCGCGAAG 454
OY	467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCTCTGCACTACTGCCACAGAACGGATCG 526
Db	455 CTAGGCAATTTCTCCGAGAGATCGTGTCTGCTCTGCACTACTGCCATGAGAACAGATGG 514
OY	527 TTCAACCGAGATCTCAAGCTGAAAAACATCTTTAGATGATGCAATGAAAACATCAAGATTG 586
Db	515 TTCACCGAGATCTCAAGCTGAAAAACATCTTTGAGATGCAATGGAATATCAAGNNNN 574
OY	587 CTGACTTTGGGCTCTCCAACCTGTACCAACAAGGCAATTCTCCAAGCGTTTCTGTGGGA 646
Db	575 NN 634
OY	647 GCCCTCTCTACGCTCGCTGAGATGATCAACGAGGAAGCCTATGTGGGCCCAAGAGTGG 706
Db	635 NNN 694
OY	707 ACAAGTGTCTCTGGGCGTCTCTCTTCAATCTCTGTGATGACCAATGCCCTTTGACG 766

Db 695 AAGAGTGTCTCCGTGGGTTCTCTCTACATCTGTGTGATGAGACCAATGCTTTTATG 754
QY 767 GGGAGATCATTAACACTGGTGAAGCAATCACTAAAGGGCTTACCTGAGCGCCCA 826
Db 755 GGGATGACCAATAGATCTAGTGAACAAGATCAACAAGGGGCTTACCGGGAGCCACTA 814
QY 827 AGCGTCGAGATGCTGTGGCTGTATCCGGGTGCTTTAATGTGTAAACCCCTGTGG 886
Db 815 AACCTCTGATGCTGTGGCTGTATCCGGGTGCTTTAATGTGTAAACCCCTGTGG 874
QY 887 CCAACTGAGAGATGTAGCCAGTCACTTGTGGTCAACTGGGGTTACACCCGAGATG 946
Db 875 CCAACCTGAGAGATGTGGCCAGTCACTGTGGGTCAACTGGGGTTACGCCACCGATGG 934
QY 947 GGGAAACGAAAGCCTGCGTGAAGGGTGGGCAACCTAGTGTGACTTTGGCCGGCTCA 1006
Db 935 GAGAGCAGAGAGCTCCGATAGAGGGTGGGCAACCTGAGAGTGACTTGGCCGGCTCA 994
QY 1007 TGGGGAATGTTAATGCTGCTCCCTCGGCGCCCTCTCTGAGAAATGAGGCCAAGTTC 1066
Db 995 TGGCTGATGCTGCTCGGCTGCTCCGCGCCCTCTCTGAGAAATGAGGCCAAGTTC 1054
QY 1067 GCTTCTTCAAGCAGCAGTGGCGGGAGGTGAAGCACTGTACTGTGGCTGGAGCGGCA 1126
Db 1055 GCTTCTTCAAGCAGCAGTGGCGGGAGGTGAAGCAACCTGTGGCTGGAGCGGCA 1114
QY 1127 ATTCTCTTAAGAGTCCGAAAGAGATGATGCTCAAAATCTGCAAGTGAACCGG 1186
Db 1115 ATTCTCTTAAGAGTCCGCAAGAGATGATGCTCAAGTGAACCGGCTCAAGTGAACCGG 1174
QY 1187 CTGAGATTAATCTTCTTCTGCGCTCGGCAAGAGCCTTAAAGCTTCCGAAAGGCTTCT 1246
Db 1175 CTGATGACATGCTCCATATGCTCTGCGCAAGAGCACTCAAGCTTCCGAAAGGCTTCT 1234
QY 1247 AGAAAGAGTCTCTACCTCGTCAAGGGAGGTGAAGAGAGCCTCAGAGAACTGACAGCG 1306
Db 1235 AGAAAGAGTCTCTACCTCGTCAAGGGAGGTGAAGAGAGCCTTCCGAGCTGAGCCCA 1294
QY 1307 TGCCTGATCTCAAGGGAGCTGTCTCTGTATCTCTGCTTCCCAAGAAAGGCTTCC 1366
Db 1295 TCCCTGAGAGCCGAGGGAGGCTGCTCC-----CCCTGCTCCCAAGAAAGGCTTCC 1345
QY 1367 TTAAGAGTCTCAAGAGCTGATCTGTGTTACTCTCTCCAGAGCCCAAGCTGCTG 1426
Db 1346 TCAAGAGGCCCAAGAGCTGCTGCTCTCTCTCCGAGCCCAAGTGAATCTG 1405
QY 1427 GGAAGCTCTTGAAGAGCTGATGTGTGATGAGTGGGAGCCCGTGAAGAGAGCTGCT 1486
Db 1406 GGAAGCTCTTGAAGAGCTGATGTGTGATGAGTGGGAGTCCCAAGAGAGAGAGCTGCT 1465
QY 1487 CACAGGCTTCAAGGGCTCTCTCTCAAGCGCAAGGAGCTTCAAACTCAATGCAAGTCT 1546
Db 1466 CCAAGGCTTCAAGGGCTCTCTCTCAAGCGCAAGGAGCTTCAAACTCAATGCAAGTCT 1525
QY 1547 CCGGCAAGCTTGAAGAGCTACCTCTGAGCACTTTGGTCTCTGAGCAACTGGCT 1606
Db 1526 CCGGCAAGCTTGAAGAGCTACCTCTGAGCACTTTGGTCTCTGAGCAACTGGCT 1585
QY 1607 CCGGCAAGCTTGAAGAGCTACCTCTGAGCACTTTGGTCTCTGAGCAACTGGCT 1666
Db 1586 CCGGCAAGCTTGAAGAGCTACCTCTGAGCACTTTGGTCTCTGAGCAACTGGCT 1645
QY 1667 TGTCTCTGAGTCTTGAAGCAATGAGCTTGTGAGCACTTTCCGAAACCCCACTGA 1726
Db 1646 TGTCTCTGAGTCTTGAAGCAAGTGTGAGCACTTTGAGCAAGTGTGAGCAAGTGTG 1705
QY 1727 GGGGCTGTGTGTGTGTGAACAACCTGAAGGGGCTTGAAGAGCTTCCCTCAAG- 1780
Db 1706 GGGGCTGTGTGTGTGTGAACAACCTGAAGGGGCTTGAAGAGCTTCCCTCAAGAGGCTG 1765
QY 1781 -----GTGAGAGCAATGTGTGAAGAACTTTGGGGAGTACTGCTTTTCTGAG 1834

Db 1766 GAAGTGTCTGAGGCGCTGGCGGAGAGATCTTTGGGGGAGAGTCTTTCCCTGAG 1825
QY 1835 ACTGCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1893
Db 1826 ACTGCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1884
RESULT 7
AY399040 1887 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION
Pan troglodytes HCM0091 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY399040
VERSION
AY399040.1 GI:39755029
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..1887
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1887
/locus_tag="HCM0091"
ORIGIN
Query Match 50.1%; Score 948.4; DB 9; Length 1887;
Best Local Similarity 60.0%; Pred. No. 2.2e-236;
Matches 1116; Conservative 0; Mismatches 722; Indels 21; Gaps 2;
QY 47 CTTCCGCTGCTGCTGAGAGCGCCGCGCTGAGCGAGCTCATCAAGTCCCTA 106
Db 35 CCACTCTCTGCGCTGAGAGCTAGCCGCGCTGAGCGAGGCTGATTAAGTCCCA 94
QY 107 AACCTTGAAG 166
Db 95 AGCCCTTAATGAAG 154
QY 167 GCTACAGATTCCTGAG 226
Db 155 GCTACAGATTCCTGAG 214
QY 227 AGAGCTCGGGGCTGCTGAG 286
Db 215 AGAGCTCGGGGCTGCTGAG 274
QY 287 AGAGCTGCTGCAATACAG 346
Db 275 AAGATGATGACATACAG 334
QY 347 TCAATGCAATCAATGAAGTGTGAAGAAATGACAGCAAGATGTGATGATGATGATGATG 406

[illegible]

OY		1487	CACAGGCTTGAAGGGCTCCTCCOTCCAACCGAAGGCATTTCAAACTCAAATGCCAAGTTCT	1546
Dd		1466	NNN	1525
OY		1547	CCCgcAcAgcctTAAGAAGCACTACCccCTTAGcACTTTggCcTCcTGgAACCATcgccCT	1606
Dd		1526	NNNNnAcAgccttgTAGAcTcTgcgcccccCACACAacCTTccgCTcCTTGTGAtGAActgcCCc	1585
OY		1607	cCTccCAatcCTtGCAGccccgcCCAGccccCTtcTcAGggggCTgtGAGtGAGAGACAGcATTcc	1666
Dd		1586	CATcTccCCCCCTTGccCCcCGggCCAGAccCTcTCAggggcTgtGAGGAGAGACAGCAcTTCc	1645
OY		1667	TGTCcTCCAGATccCTttTGACCAATTTgACTTgcCTGAAGcCTTCcCCGAACccCACTGA	1726
Dd		1646	TGTCcTCTGMAcTcCTTTtGACCAcGTGAactTtgCTtGAACccGCTccCAgAcCCcCAcTgc	1705
OY		1727	GggGcTgTgtGTcTGTgSGACAAcCTGAgGggGgCTTgAGAcCcTCCcCTCAGAAG-----	1780
Dd		1706	GggGcTgTgtGTcTGTgSACAacCTGAAtGggGgCTTgAGAGccCCcCTcAGAAGgCcCTg	1765
OY		1781	-----GTCTGAACCAATGtGTgSGAGAAATcTTggGGGGAATAgCTGTTTTCTCTGACAG	1834
Dd		1766	GmAcTGCCTGTGAGGcGCTgGcGcGAGAGATcCTTYtGGggGACAGCTGCTYYTCCCTGACAG	1825
OY		1835	ACTGcCAAGAGGTGACTGCAGCCCAAGACAGACAAAGcCCTAGGAATCTGCTCAAAAGCTcAGC	1893
Dd		1826	ACTcGcAGAGAGtGACAGcCAcCTACcCGGcAGCACTGAaggGtCTGCTCAAAAGCTcAACC	1884
RESULT 8				
CR749209		7004 bp	mRNA	linear HTC 19-AUG-2004
LOCUS				
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686F01113 (from clone DKFZp686F01113).			
ACCESSION	CR749209			
VERSION	CR749209.1 GI:51476151			
KEYWORDS	HTC.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homniidae; Homo.			
	1 (bases 1 to 7004)			
	Otenwaelder,B., Obermaier,B., Deutschenbauer,S., Schaiipp,A.,			
	Mewes,H.W., Well,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and			
	Wiemann,S.			
	The German CDNA Consortium			
CONSRPTM	Direct Submission			
TITLE	Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764			
JOURNAL	Nuehnberg, GERMANY			
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp686F01113) is available at the RZPD Deutsches Ressourcentzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686F01113 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ . Location/Qualifiers 1..7004 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp686F01113" /tissue_type="fetal kidney" /clone_id="686 (synonym: nlcc3). Vector psportL_Sfi; host DHIOB; sites SfilA + StIIIB" /dev_stage="fetal" /note="hypothetical protein, N-terminus truncated, not fully spliced" 1..7004 /gene="DKFZp686F01113"			
FEATURES				
SOURCE				
gene				

CDS

<4163..5479
/gene="Dkx2p66f01113"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18066.1"
/db_xref="GI:51476152"
/translation="IADPLSNLYHOKKPIQTPGSPLYASPEIUNKPYTPREVDSW
SLVILYIIVHGTMPDPGHDKIIVKQISNAYREPPSPDACLIMLWNPTRRA
TLEDVASHMWNWGYATRVGEAPHEGHPGSDSASAMADWLRSRPLENGAV
CSFPKQAPGGSTTGLERQSHLKKRKENDMAQSHSDPADTARPKSNLKP
GILKKVYASABGVQEDPPELSPAPSGQAPLPLPKGILKKRQESGVSSPEP
ESGELDAGAVFVQDEPKQSPQASGLLHKRQILKNGKFSOTLLELAPTFGL
DELAPPLARASPSGAVSBDSTLSSESFDLPLRLEPPRGCVSDNLTGLEB
PSPBGRPSCLRRKQDPLGDSCSFULTQGEVATYRQALRVCSKLT"

Query Match 48.9%; Score 925.4; DB 3; Length 7004;
Best Local Similarity 82.2%; Pred. No. 3,2e-230;
Matches 1096; Conservative 0; Mismatches 216; Indels 21; Gaps 2;

573 AAACATCAAGATTGCTGACTTTGGCTCTTCAACCTGTACCAAAAGCAAGTTCTTCA 632
4153 AATCTGACAGATTGCTGACTTGGCTCTTCAACCTGTACCAAGCAAGTTCTTCA 4212
633 GAGCTTGTGGAGAGCCCTCTCTACGCTGCGCTGAGATAGTCAACGGAAAGCCCTATGT 692
4213 GACATTTGTGGAGAGCCCTCTCTACGCTGCGCTGAGATAGTCAACGGAAAGCCCTATGT 4272
693 GGGCCCAAGAGTGAAGAGTGTCTGTGGCGTCTCTGTACATCTGTGTGATGAC 752
4273 AGGCCCAAGAGTGAAGAGTGTCTGTGGCGTCTCTGTACATCTGTGTGATGAC 4332
753 CATGCTTGTGAAGAGAGATCATATAACCTGTGTGAAGCAATCAAGTAAAGGGGCTTGA 812
4333 CATGCTTGTGAAGAGAGATCATATAACCTGTGTGAAGCAATCAAGTAAAGGGGCTTGA 4392
813 CCGTAGCCGCGCCCAAGCCGTCCGATGCTGTGGCTGATCCGGTGGCTTAAAGTGGA 872
4393 CCGTAGCCGCGCCCAAGCCGTCCGATGCTGTGGCTGATCCGGTGGCTTAAAGTGGA 4452
873 CCGTAGCCGCGCCCAAGCCGTCCGATGCTGTGGCTGATCCGGTGGCTTAAAGTGGA 932
4453 CCGTAGCCGCGCCCAAGCCGTCCGATGCTGTGGCTGATCCGGTGGCTTAAAGTGGA 4512
933 CACCAACGAGTGGAGGAGACGAAAGCCCTGTGAGGTGGGCAACCTAGTGTGACTT 992
4513 CCGCACCCGAGTGGAGGAGACGAGGCTCCGATGAGGTGGGCAACCTAGTGTGACTT 4572
993 TGGCCGGGCTCCATGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1052
4573 TGGCCGGGCTCCATGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4632
1053 AGCCAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1112
4633 GGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4692
1113 GTTGAAGCGGCAATTTCTTAAAGATCCCGAAGAGATGACATGGCTCAAAATCT 1172
4693 CTTGAGCGGCAATTTCTTAAAGATCCCGAAGAGATGACATGGCTCAAAATCT 4752
1173 GCAAGTGAAGCGGCTGAGATGATCTTCTGCGCTGCGCAAGAGCGCTTAAAGTCTTCC 1232
4753 CCAAGTGAAGCGGCTGAGATGATCTTCTGCGCTGCGCAAGAGCGCTTAAAGTCTTCC 4812
1233 GAAAGCATCTCAAGAAAAGTCTTACCTGTCAGGGAGGTATACAGAGAGAGCCCTCA 1292
4813 AAGAGCATCTCAAGAAAAGTCTTACCTGTCAGGGAGGTATACAGAGAGAGCCCTCA 4872
1293 GGAATCTCAAGCGGTGCTGATATCTCAAGGAGAGCTGTCTCTGCTGATCTCTCTCC 1352
4873 GGAATCTCAAGCGGTGCTGATATCTCAAGGAGAGCTGTCTCTGCTGATCTCTCTCC 4923
1353 AAGAAAAGGATCTTAAAGATCTGACAGCGGATCTGGTTACTTACTCTCTCCAGA 1412

Db 4924 CAAAGAGGAGATCTTCAAGAGAGCCCGACAGCGGAGTCTGGACTACTTCTTCCGA 4993
1413 GCCCAGGAGTCTGGAGAACTTGAAGAGCAGATGATGTGTTGTGAGTGGAGAGCCCGT 1472
4984 GCCCAGGAGTCTGGAGAGCTTGAAGAGCAGAGGAGATGTGTTGTGAGTGGAGATCCCA 5043
1473 GAGCAGAGTCTTCAAGAGCTTGAAGAGCTTCTTCTTCAACCGAGAGGCAATTCGAAC 1532
5044 GAGCAGAGTCTTCAAGAGCTTGAAGAGCTTCTTCTTCAACCGAGAGGCAATTCGAAC 5103
1533 CAATGGAAGTCTTCCCGCAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1592
5104 CAATGGAAGTCTTCCCGCAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5163
1593 GAGCAGAGTCTTCCCGCAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1652
5164 GAGCAGAGTCTTCCCGCAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5223
1653 TGAAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1712
5224 GAGCAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5283
1713 CGAAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1772
5284 AGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5343
1773 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1820
5344 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5403
1821 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1880
5404 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 5463
1881 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1940
5464 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1997

Db 1881 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1940
5464 CTGAGAGAGTCTTCAAGAGCTTGAAGAGCTTCAACCGAGAGGCAATTCGAAC 1997

RESULT 9
CB248251
LOCUS
DEFINITION
UT-M-FD0-bv1-a-21-0-UT.r1 NIH BMAP_F00 Mus musculus cDNA clone
IMAGE:5718428 5', mRNA sequence.
CB248251
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
REFERENCE
AUTHORS
TITLES
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
CDNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP).
Seq primer: pYX-5.
Location/Qualifiers
1..726
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"

BC174288 805 bp mRNA linear EST 06-FEB-2001
LOCUS BC174288
DEFINITION 602334589P1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457398 5',
mRNA sequence.
ACCESSION BC174288
VERSION BC174288.1 GI:12680991
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 805)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM10254 row: b column: 23
High quality sequence stop: 704.
location/Qualifiers
1..805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457398"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/clone_1lb="NCI CGAP Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

Query Match 34.4%; Score 651.6; DB 4; Length 805;
Best Local Similarity 96.6%; Pred. No. 7.1e-159;
Matches 709; Conservative 0; Mismatches 19; Indels 6; Gaps 4;

QY 379 AGCAAGTTGATGTCATGAGTATGCGAGCGGCGATCTGTATGATTACATCACT 438
1 AGCAAGTTGATGTCATGAGTATGCGAGCGGCGATCTGTATGATTACATCACT 60
QY 439 GAGGCGCAGCGGCTGATGAGCGGAGCGGCGGCGATTTCTTCCAGCATGTGTCTGCC 498
61 GAGGCGCAGCGGCTGATGAGCGGAGCGGCGGCGGCGATTTCTTCCAGCATGTGTCTGCC 120
QY CTGCACTACTGCGCAGCAAGAGGAGTGTTCACCGAGATCTCAAGCTGGAAACATCCTT 558
121 CTGCACTACTGCGCAGCAAGAGGAGTGTTCACCGAGATCTCAAGCTGGAAACATCCTT 180
QY CTAGATGCCAATGAAACATCAAGATTCTGACTTTGGCCCTTCCAACTGTACACAAA 618
181 CTAGATGCCAATGAAACATCAAGATTCTGACTTTGGCCCTTCCAACTGTACACAAA 240
QY 619 GGCAGATTCTCCAGAGATTCTGTGGGAGAGCCCTCTTACGCGCTGCTGAGATGTCAAC 678
241 GGCAGATTCTCCAGAGATTCTGTGGGAGAGCCCTCTTACGCGCTGCTGAGATGTCAAC 300
QY GGAAGCCCTATGTGGGCGGAGAGTGAAGAGTGTCTGTGGGCGTTCTCTGTACATC 738
679 GGAAGCCCTATGTGGGCGGAGAGTGAAGAGTGTCTGTGGGCGTTCTCTGTACATC 738
Db 301 GGAAGCCCTATGTGGGCGGAGAGTGAAGAGTGTCTGTGGGCGTTCTCTGTACATC 360
QY CTGTGTGATGCGACCATGCGCTTTGACGGGCGAGATCAATAAACACTGGTGAACCAATC 798
739 CTGTGTGATGCGACCATGCGCTTTGACGGGCGAGATCAATAAACACTGGTGAACCAATC 798

Db 361 CTGTGTGATGCGACCATGCGCTTTGACGGGCGAGATCAATAAACACTGGTGAACCAATC 420
QY 799 AGTAACGGGCGCTTACCGTGAAGCGCGCCAGCCGCTCCGATGCTGTGGCTGATCCGTTGG 858
421 AGTAACGGGCGCTTACCGTGAAGCGCGCCAGCCGCTCCGATGCTGTGGCTGATCCGTTGG 480
QY 859 CTGTTAATGTGAACCCCAACCCCGTGGGCGCACTGGAAGATGTAGGCACTGTTGGTGG 918
481 CTGTTAATGTGAACCCCAACCCCGTGGGCGCACTGGAAGATGTAGGCACTGTTGGTGG 540
QY 919 GTCAACTGGGGTTACACACCGAGATGTGGGGAACAGAAAGCCCGTGAAGGGTGGCAC 978
541 GTCAACTGGGGTTACACACCGAGATGTGGGGAACAGAAAGCCCGTGAAGGGTGGCAC 600
QY 979 CCTAGTGTACTTTGGCGCGGCTCCATGAGCGGAGCTGTACGTCGCTCCGCGCC 1038
601 CCTAGTGTACTTTGGC--GGGTCATATGGCGGACGTGACTGCTC--TGGCGCC 656
QY 1039 CTCTGGAGATGAGGCGCAAGTGTGACGCTTTCAAGACAGACGTCGCGGAGTGA 1098
657 CTCTGGAGATGAGGCGCAAGTGTGACGCTTTCAAGACAGACGTCGCGGAGTGA 714
QY 1099 AGCACTGTACTGG 1112
Db 715 AGCACTGTACTGG 728

RESULT 13
BC694881 646 bp mRNA linear EST 04-MAY-2001
LOCUS BC694881
DEFINITION NISC_1v09h10.w1 Soares NMBP2 pituitary Mus musculus cDNA clone
IMAGE:4317546 5', mRNA sequence.
ACCESSION BC694881
VERSION BC694881.1 GI:13954237
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 646)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium/LNLN
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
MGI:1598314
Plate: LLM9921 row: 0 column: 19
Seq primer: T7 primer.
location/Qualifiers
1..646
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4317546"
/tissue_type="pituitary gland"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="Soares NMBP2 pituitary"
/note="Organ: brain; Vector: pT73D-Pac; Site 1: NotI;
Site 2: EcoRI; 1st strand cDNA was primed with a NotI -
oligo(dT) primer
5'-ACTGGAAGAATTGCGCGCGCGCGCTTTTCTTTTCTTTT-3';
double-stranded cDNA was ligated to EcoRI adaptors
5'-ATTTCGCGACGAG-3' AND 5'-CTTCGCGG-3' (Pharmacia),
digested with NotI and cloned into the NotI and EcoRI

ORIGIN

sites of the pT73D-Pact vector. Library went through one round of normalization, and was constructed in the laboratory of M. Bento Soares (University of Iowa)."

Query Match

34.1%; Score 646; DB 4; Length 646;

Best Local Similarity 100.0%; Pred. No. 1,9e-157; Indels 0; Gaps 0;

Matches 646; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES

source

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
plate: LHAM811 row: 9 column: 12
High quality sequence stop: 651.
Location/Qualifiers
1. 948
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4223267"
/lab_host="DH10B (T1 phage-resistant)"
/clone_11b="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NCI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match

34.1%; Score 645.8; DB 2; Length 948;

Best Local Similarity 96.3%; Pred. No. 2.4e-157; Indels 4; Gaps 2;

Matches 683; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

QY 1193 ATACCTCTTCTGCGCCCTGGCAGAGCAGCCTTAAGCTTCCGAAAGGCAATTCACAGAAA 1252
DB 1 ATACCTCTTCTGCGCCCTGGCAGAGCAGCCTTAAGCTTCCGAAAGGCAATTCACAGAAA 60
QY 1253 AGTCTCTTACTCTGCTCAGGGAGAGTACAGAGAGACCTCTGAGAACTCAGACCGGTGCTG 1312
DB 61 AGTCTCTTACTCTGCTCAGGGAGAGTACAGAGAGACCTCTGAGAACTCAGACCGGTGCTG 120
QY 1313 ATACTCCAGGGAGCAGCCTGCTGCTGATCCCTGCTCCAGAGAAAGGCAATTCCTTAA 1372
DB 121 ATACTCCAGGGAGCAGCCTGCTGCTGATCCCTGCTCCAGAGAAAGGCAATTCCTTAA 180
QY 1373 AGTCTCCAGAGCAGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1432
DB 181 AGTCTCCAGAGCAGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 1433 TCTTGAAGCAGAGTGAATGTTTGTGAGTGGGAGCCCGGTGAGAGAGAGTCTTCAAG 1492
DB 241 TCTTGAAGCAGAGTGAATGTTTGTGAGTGGGAGCCCGGTGAGAGAGAGTCTTCAAG 300
QY 1493 CTTCAGAGGCTCTCTCTCCAGCAGAGGCAATTCCTGCTGCTGCTGCTGCTGCTGCTG 1552
DB 301 CTTCAGAGGCTCTCTCTCCAGCAGAGGCAATTCCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 1553 CAGCCTTGAAGAGCACTACCCCTGAGCACTTGGCTCCCTGAGCACTGAGCTCTGCT 1612
DB 361 CAGCCTTGAAGAGCACTACCCCTGAGCACTTGGCTCCCTGAGCACTGAGCTCTGCT 420
QY 1613 ATCTGAGAGCCCGGAGCCAGCCGCTCAGAGGCTGAGAGTGAAGAGAGAGTCTGCT 1672
DB 421 ATCTGAGAGCCCGGAGCCAGCCGCTCAGAGGCTGAGAGTGAAGAGAGAGTCTGCT 480
QY 1673 CCGAGTCTTGAAGCAATTTGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1732
DB 481 CCGAGTCTTGAAGCAATTTGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 1733 GTGTGTGTGAGCAACCTGAGAGGAGCTTGAAGAGCTCTGCTGAGAGAGTCTGAG 1792
DB 541 GTGTGTGTGAGCAACCTGAGAGGAGCTTGAAGAGCTCTGCTGAGAGAGTCTGAG 600
QY 1793 GGTGGCAGGAATCTTGGGAGATGCTGCTTCTGAGCACTG 1838
DB 601 GGTGGCAGGAATCTTGGGAGATGCTGCTTCTGAGCACTG 646

RESULT 14

LOCUS

BF789245 948 bp mRNA linear EST 12-JAN-2001

DEFINITION 602105144F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223267

5', mRNA sequence.

ACCESSION

BF789245

VERSION

BF789245.1

KEYWORDS

GI:12094281

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

NIH-MGC

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

QY 867 GGTGAACCCACCGCTGCGGACACACTGAGAGTGAAGCAGTCAATTTGGTGTCACTG 926
DB 12 GGTGAACCCACCGCTGCGGACACACTGAGAGTGAAGCAGTCAATTTGGTGTCACTG 71
QY 927 GGGTTACACCAACCGAGTGGGAGACAGAAAGCTTGGCTGAGAGGTGGACCTTAAG 986
DB 72 GGGTTACACCAACCGAGTGGGAGACAGAAAGCTTGGCTGAGAGGTGGACCTTAAG 131
QY 987 TGACTTTGGCGGCGCTCCATGAGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1046
DB 132 TGACTTTGGCGGCGCTCCATGAGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 191
QY 1047 GAATGAGCCCAAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1106
DB 192 GAATGAGCCCAAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 251
QY 1107 ACCTGAGCTGAGAGCGGCAATTTCTTAAAGAGTCCGAAAGAGAGATGAGTCA 1166
DB 252 ACCTGAGCTGAGAGCGGCAATTTCTTAAAGAGTCCGAAAGAGAGATGAGTCA 311
QY 1167 AATCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1226
DB 312 AATCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 371
QY 1227 GCTTCGAAAGGCAATTTCTCAAGAAAGTCTTACCTGCTGAGGAGGTAACAGAGGA 1286
DB 372 GCTTCGAAAGGCAATTTCTCAAGAAAGTCTTACCTGCTGAGGAGGTAACAGAGGA 431
QY 1287 CCCTCAGAACTCAGACCGGTGCTGATCTCAGGAGCAGCTGCTGCTGATCTCT 1346
DB 432 CCCTCAGAACTCAGACCGGTGCTGATCTCAGGAGCAGCTGCTGCTGATCTCT 491
QY 1347 GCTTCGAAAGGCAATTTCTTAAAGTCTGAGAGGCTGATCTGATCTGATCTCT 1406
DB 492 GCTTCGAAAGGCAATTTCTTAAAGTCTGAGAGGCTGATCTGATCTGATCTCT 551
QY 1407 TCCAGAGCCACGAGTCTGAGGAGCTTAAAGAGCAGCAGTGAAGTGTGAGTGAAGGGA 1466
DB 552 TCCAGAGCCACGAGTCTGAGGAGCTTAAAGAGCAGCAGTGAAGTGTGAGTGAAGGGA 611
QY 1467 CCCCGTGAAGCAGAGTCTTCAAGGCTTCAAGGCTCTCTTCAAGCAGCAGCAGTCT 1526
DB 612 CCCCGTGAAGCAGAGTCTTCAAGGCTTCAAGGCTCTCTTCAAGCAGCAGCAGTCT 669
QY 1527 CAACTCAATGGCAAGTCTCCCGCAGAGCTTAAAGAGGCACTACCTT 1575

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:12:18 ; Search time 162.033 Seconds
(without alignments)
1994.176 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293
Sequence: 1 MESVALLQRPSPQAPSASALA.....DCQEVTAAYRQALGICSKLS 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trernbl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	2 Q9DBV0	Q9DBV0 m mus muscu
2	3285	99.8	631	2 Q8CIC0	Q8CIC0 mus musculu
3	3279	99.6	639	2 Q80ZM3	Q80ZM3 mus musculu
4	3272	99.4	639	2 Q8BZM4	Q8BZM4 mus musculu
5	3087.5	93.8	630	2 Q66HES	Q66HES rattus norv
6	2824.5	85.8	638	2 Q9H093	Q9H093 homo sapien
7	1949.5	59.2	438	2 Q68E04	Q68E04 homo sapien
8	1699.5	51.6	661	1 ARKS HUMAN	Q60285 homo sapien
9	1695.5	51.5	658	2 Q641K5	Q641K5 mus musculu
10	1515.5	46.0	575	2 Q616D6	Q616D6 mus musculu
11	1008.5	30.6	461	2 Q8CGE1	Q8CGE1 mus musculu
12	888	27.0	1551	2 Q65ZM4	Q65ZM4 caenorhabdi
13	888	27.0	1592	2 Q65ZM3	Q65ZM3 caenorhabdi
14	881.5	26.8	1180	2 Q7KSS0	Q7KSS0 drosophila
15	881.5	26.8	1180	2 Q7YU19	Q7YU19 drosophila
16	881.5	26.8	1427	2 Q9VH05	Q9VH05 drosophila
17	853	25.9	383	2 Q7QBH4	Q7QBH4 anopheles g
18	748	22.7	752	1 MRK4 HUMAN	Q96104 mus musculu
19	745	22.6	752	1 Q8CIP4	Q8CIP4 mus musculu
20	743	22.6	752	2 Q8NG37	Q8NG37 homo sapien
21	716.5	21.8	792	2 Q6TNT7	Q6TNT7 xenopus lae
22	701	21.3	729	2 Q9JKE4	Q9JKE4 mus musculu
23	701	21.3	744	2 Q9VKE5	Q9VKE5 mus musculu
24	700.5	21.3	797	2 Q8VHF0	Q8VHF0 rattus norv
25	696.5	21.1	780	2 Q804T1	Q804T1 xenopus lae
26	695.5	21.1	993	2 Q7KXK4	Q7KXK4 drosophila
27	693.5	21.1	795	2 Q8VHJ5	Q8VHJ5 mus musculu
28	693	21.0	778	2 Q96HB3	Q96HB3 homo sapien
29	692	21.0	725	2 Q804T2	Q804T2 xenopus lae
30	691.5	21.0	793	2 Q08678	Q08678 rattus norv
31	691.5	21.0	795	2 Q9P0L2	Q9P0L2 homo sapien

32	691	21.0	712	2 Q7QBH3	Q7QBH3 anopheles g
33	690.5	21.0	755	2 Q7KZ17	Q7KZ17 homo sapien
34	690.5	21.0	1098	2 Q7KXK7	Q7KXK7 drosophila
35	688.5	20.9	888	2 Q8BR95	Q8BR95 mus musculu
36	688	20.9	745	2 Q15524	Q15524 homo sapien
37	686.5	20.8	774	1 MRK2 MOUSE	Q05512 mus musculu
38	685.5	20.8	705	2 Q9VGF9	Q9VGF9 drosophila
39	684.5	20.8	771	2 Q69Z17	Q69Z17 mus musculu
40	684	20.8	832	2 Q963B6	Q963B6 drosophila
41	680.5	20.7	719	2 Q68A18	Q68A18 homo sapien
42	680	20.6	776	2 Q7ZYL7	Q7ZYL7 xenopus lae
43	680	20.6	785	2 Q8GCV3	Q8GCV3 xenopus lae
44	678.5	20.6	691	2 Q96RG0	Q96RG0 homo sapien
45	678.5	20.6	834	2 Q8MVX1	Q8MVX1 haemochus

ALIGNMENTS

RESULT 1	Q9DBV0	PRELIMINARY;	PRT;	631 AA.
AC	Q9DBV0:			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013822 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) (Mus musculus adult male cecum			
DE	CDNA, RIKEN full-length enriched library, clone:9130215K18			
DE	product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537).			
GN	Name=1200013822R1K;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;			
RA	Carninci P., Hayashizaki Y.;			
RT	"High-efficiency full-length cDNA cloning.";			
RL	Meth. Enzymol. 303:19-44 (1999).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;			
RA	RIKEN FANTOM Consortium;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690 (2001).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RA	The FANTOM Consortium,			
RT	"The RIKEN Genome Exploration Research Group Phase I & II Team;			
RL	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";			
RL	Nature 420:563-573 (2002).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;			
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,			
RT	Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;			
RL	"Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";			
RL	Genome Res. 10:1617-1630 (2000).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;			
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,			

RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto S., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.",
RL Genome Res. 10:1757-1771(2000).

RC [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiroaki T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsumura T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takasashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaki T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kaich H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsumura T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saesaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AK036737; BAB23518.1; -
DR HSSP; P31751; IGZK.
DR MGD; MGI:1921387; 1200013B22Rik.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002230; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR Pfam; PF00069; PKinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 631 AA; 69806 MW; D51C042DE6C174B CRC64;

Query Match 100.0%; Score 3293; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 6e-181;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLRPSQAPASALASASAPLADGLTKSPKLMKKQAVRRHHKHLRRHYEFL 60
DB 1 MESVALLRPSQAPASALASASAPLADGLTKSPKLMKKQAVRRHHKHLRRHYEFL 60
QY 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIKDEQDLHIREIEMSLNHPHIIAH 120
DB 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIKDEQDLHIREIEMSLNHPHIIAH 120

QY 121 EYFENSSKVIIVMEYASRGDLVDYTSERPRLSERPARRHFRQIVSNLHYCHONGIYHRL 180
DB 121 EYFENSSKVIIVMEYASRGDLVDYTSERPRLSERPARRHFRQIVSNLHYCHONGIYHRL 180
QY 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASBEIYNGKRYVPEVDSMSL 240
DB 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASBEIYNGKRYVPEVDSMSL 240
QY 241 GVLVLYLVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLMVNPTRATLED 300
DB 241 GVLVLYLVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLMVNPTRATLED 300
QY 301 VASHMMVMVNGYTTGCGEAGALREGGHPSGDFERRASNAWDLRRSSRPLLENGAKVCSFFQ 360
DB 301 VASHMMVMVNGYTTGCGEAGALREGGHPSGDFERRASNAWDLRRSSRPLLENGAKVCSFFQ 360
QY 361 HYPGGGSLVPGLEPRLSLKSKRENDMAONLGDPAEDTSSRPGKSLTLPGIILKKSS 420
DB 361 HYPGGGSLVPGLEPRLSLKSKRENDMAONLGDPAEDTSSRPGKSLTLPGIILKKSS 420
QY 421 TSSGEVQEDPQELRPVPTDPPQPVPAVSLPRKGIILKKSRORESGYSSPEPSEGEILD 480
DB 421 TSSGEVQEDPQELRPVPTDPPQPVPAVSLPRKGIILKKSRORESGYSSPEPSEGEILD 480
QY 481 ASDVFSVSGDPVPSKSPQASGLLHKKGIILKNGKSRPLLEGTPTSPSLDQLASSHPA 540
DB 481 ASDVFSVSGDPVPSKSPQASGLLHKKGIILKNGKSRPLLEGTPTSPSLDQLASSHPA 540
QY 541 ARPSRPSGAVSDSLTSSSEFPDQLPERLPEPTPRGCVSVNLTGLEOPSEGLKRWQ 600
DB 541 ARPSRPSGAVSDSLTSSSEFPDQLPERLPEPTPRGCVSVNLTGLEOPSEGLKRWQ 600
QY 601 ESLGDSCESLTDCQEVTAAYRQALGICSKTSS 631
DB 601 ESLGDSCESLTDCQEVTAAYRQALGICSKTSS 631

RESULT 2
Q8CIC0 PRELIMINARY; PRT; 631 AA.
AC Q8CIC0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
GN 1200013B22Rik protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strusberg R.D., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.J., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzyzanski M.I., Skalski U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Mair M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 CC Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC033302; AAH3302.1; -.
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SMO0220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 631 AA; 69745 MW; C6C904224D9174F CRC64;

Query Match 99.8%; Score 3285; DB 2; Length 631;
 Best Local Similarity 99.8%; Pred. No. 1.7e-180;

Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 DB 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 QY 61 ETLGKCTYGVKKARSSGRLVAIKIRKDKIDQDLHRRREIEMSSLNPHITAIH 120
 DB 61 ETLGKCTYGVKKARSSGRLVAIKIRKDKIDQDLHRRREIEMSSLNPHITAIH 120
 QY 121 EVFENSSTKIYVNEVYASRGDLVYISRRPRLSRDPAHFRQVSAHYHONGIYARDL 180
 DB 121 EVFENSSTKIYVNEVYASRGDLVYISRRPRLSRDPAHFRQVSAHYHONGIYARDL 180
 QY 181 KLENIILLDANGNIKIDFGLSNTYHKGFQTFGSPLYASPEIYVNGKPYVGEVDSMSL 240
 DB 181 KLENIILLDANGNIKIDFGLSNTYHKGFQTFGSPLYASPEIYVNGKPYVGEVDSMSL 240
 QY 241 GVLTYLIVHGTMFPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 DB 241 GVLTYLIVHGTMFPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 QY 301 VASHMMVWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRRLLENGAKVCSFQK 360
 DB 301 VASHMMVWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRRLLENGAKVCSFQK 360
 QY 361 HVFGGGSSTVPEGLRQHSLLKSRKENDMAQNLQGPADTSRRPKSSLKLPKGIKKKSS 420
 DB 361 HVFGGGSSTVPEGLRQHSLLKSRKENDMAQNLQGPADTSRRPKSSLKLPKGIKKKSS 420
 QY 421 TSSGEVQEDQELRPVDPDPQVPVAVSLPRKGIKKSRQRESGYSSPEPSSEGLLD 480
 DB 421 TSSGEVQEDQELRPVDPDPQVPVAVSLPRKGIKKSRQRESGYSSPEPSSEGLLD 480
 QY 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALEGTTPTFGSLDOLASSHPA 540
 DB 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALEGTTPTFGSLDOLASSHPA 540
 QY 541 ARSPRSGAVSEDSIISSEFPDLDLPERLPETPLRCVGVNDLRGLQEPSEGLKRWQ 600
 DB 541 ARSPRSGAVSEDSIISSEFPDLDLPERLPETPLRCVGVNDLRGLQEPSEGLKRWQ 600
 QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
 DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 3
 ID Q80ZM3 PRELIMINARY; PRT; 639 AA.
 AC Q80ZM3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 1200013B22Rik protein.
 GN Name=1200013B22Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
 RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
 RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
 RA Hopkins R.F.; Jordan H.; Moore T.; Max S.T.; Wang J.; Heien F.;
 RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
 RA Scapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.B.;
 RA Brownstein M.J.; Uedin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
 RA Raha S.S.; Loughellano N.A.; Peters G.J.; Abramson R.D.; Mullaly S.J.;
 RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
 RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.U.; Huiyk S.W.;
 RA Valliion D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
 RA Fey J.J.; Helton E.; Kettelman M.; Madan A.; Rodriguez S.; Sanchez A.;
 RA Whiting R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
 RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Buttefield Y.S.;
 RA Krzywiński M.I.; Skalka U.; Smalins D.E.; Schnerch A.; Schein J.E.;
 RA Jones S.J.; Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 CC Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC046833; AAH46833.1; -.
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SMO0220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 639 AA; 70675 MW; FB9C40228F53872C CRC64;

Query Match 99.8%; Score 3279; DB 2; Length 639;
 Best Local Similarity 98.7%; Pred. No. 3.9e-180;
 Matches 631; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 DB 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60

QY 61 ETLGKGTGKVKKARSSGRLVAKSRKDKITDEOULLHREIEIEMSLNPHIAT 120
DB 61 ETLGKGTGKVKKARSSGRLVAKSRKDKITDEOULLHREIEIEMSLNPHIAT 120
QY 121 E-----VFENSSKIVIMVAYASRDLDYISERPRLSERDARHPFRQIVSAHYCHQ 172
DB 121 EVGRSLVTVFENSSKIVIMVAYASRDLDYISERPRLSERDARHPFRQIVSAHYCHQ 180
QY 173 NGIVRDLKLENTLLDANGNIKADPGLSNLHKKFLQTFCCSPLYASPEIYNGKPYG 232
DB 181 NGIVHDLKLENTLLDANGNIKADPGLSNLHKKFLQTFCCSPLYASPEIYNGKPYG 240
QY 233 PEYDSMSIGVLYILVHGTMPEFGQDHKTUKQISNAYREPPKPSACGLIIMLWMP 232
DB 241 PEYDSMSIGVLYILVHGTMPEFGQDHKTUKQISNAYREPPKPSACGLIIMLWMP 300
QY 293 TRRATLEDVASHMVMVNGYTTGVEOEDALREGGSPGDFRASMADMLRRSSRPLENGA 352
DB 301 TRRATLEDVASHMVMVNGYTTGVEOEDALREGGSPGDFRASMADMLRRSSRPLENGA 360
QY 353 KVCSPFKOHVPGGSGTVPGLEHOSLKKSRKENDMAQNTQGDPAEDTSSRPKGSLSLTPK 412
DB 361 KVCSPFKOHVPGGSGTVPGLEHOSLKKSRKENDMAQNTQGDPAEDTSSRPKGSLSLTPK 420
QY 413 GLIKKSSVSSGVEODPOLRVPPTPGQPVPAVSLPRKGLIKTSKRSQESGYSSP 472
DB 421 GLIKKSSVSSGVEODPOLRVPPTPGQPVPAVSLPRKGLIKTSKRSQESGYSSP 480
QY 473 SESGELLDAADVVSQDPVEQKSPQASGLLHRRKGLIKNGKESRTALREGTSTFGSLD 532
DB 481 SESGELLDAADVVSQDPVEQKSPQASGLLHRRKGLIKNGKESRTALREGTSTFGSLD 540
QY 533 QLAASHPARAPSPGSAVSEDSILSSESPQDLPERLPETPLRGCVSVNMLGLOPPS 592
DB 541 QLAASHPARAPSPGSAVSEDSILSSESPQDLPERLPETPLRGCVSVNMLGLOPPS 600
QY 593 EGIKRWMOESIGDSCSLTDCCEVTAAVRAQALGICSTLS 631
DB 601 EGIKRWMOESIGDSCSLTDCCEVTAAVRAQALGICSTLS 639
RESULT 4
Q8BZM4 PRELIMINARY; PRT; 639 AA.
AC Q8BZM4; PRELIMINARY; PRT; 639 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:93J015AN24 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.1-).
GN Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99379253; PubMed=10349636; DOI=10.1016/S0076-6679(99)03004-9;
RT "High-efficiency full-length cDNA cloning.";
RT Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RT RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subcloning of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kono H., Akiyama J., Nishi K., Kikunai T., Taahiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hironaka T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohnaka N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takada Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AK034082; BAC28575.1; -.
DR HSSP; P31751; IGZK.
DR MGD; MGI:1921387; 1200013B22R1k.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot. kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot. kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 639 AA; 70632 MW; FB98EE915C95FASD CRC64;

Query Match 99.4%; Score 3272; DB 2; Length 639;
Best Local Similarity 98.6%; Pred. No. 9,8e-180;
Matches 630; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
QY 1 MESVALQRPQAPASALASAPRLADGLIKSPKPLKKQAVKRRHHKHLRHYEFL 60
DB 1 MESGALQRPQAPASALASAPRLADGLIKSPKPLKKQAVKRRHHKHLRHYEFL 60

QY 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 120
DB 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 120
QY 121 E-----VFENSSKTIYVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQ 172
DB 121 EVGRSLVTVFENSSKTIYVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQ 180
QY 173 NGIVHBDLKLLENLILNANGNIKADFGLSNLHYHKGFLQTFCCSPLYASPEIYNGKPYVG 232
DB 181 NGIVHBDLKLLENLILNANGNIKADFGLSNLHYHKGFLQTFCCSPLYASPEIYNGKPYVG 240
QY 233 PEVDNSLSGLVLLYLIVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLLMVP 292
DB 241 PEVDNSLSGLVLLYLIVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLLMVP 300
QY 293 TRRATLEVDASHHWVMWGTGTVGEOBALREGHPSGDFPRASMDWLRSSRPILLENGA 352
DB 301 TRRATLEVDASHHWVMWGTGTVGEOBALREGHPSGDFPRASMDWLRSSRPILLENGA 360
QY 353 KVCSPFKOHVPGGSGTVPGLEHOSLKKSRKENDMAQNLQGDPAEDTSSRPGKSLKLPK 412
DB 361 KVCSPFKOHVPGGSGTVPGLEHOSLKKSRKENDMAQNLQGDPAEDTSSRPGKSLKLPK 420
QY 413 GILKKKSTSSGVEQEDPOLRPPVPTPGQPVPAVSLPRKGIILKKSQRSGYSSP 472
DB 421 GILKKKSTSSGVEQEDPOLRPPVPTPGQPVPAVSLPRKGIILKKSQRSGYSSP 480
QY 473 SESGELLDSADVSGDPVEQKSPQASGLILHRKGIILKNGKSRRLBETTPSTGSLD 532
DB 481 SESGELLDSADVSGDPVEQKSPQASGLILHRKGIILKNGKSRRLBETTPSTGSLD 540
QY 533 QLAASHPARPSRPSGAVSDSLSSSPDQLPERLPETPLRGCVSVNLAGLQPPS 592
DB 541 QLAASHPARPSRPSGAVSDSLSSSPDQLPERLPETPLRGCVSVNLAGLQPPS 600
QY 593 EGLKRWQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 601 EGLKRWQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 639

RESULT 5
Q66HES PRELIMINARY; PRT; 630 AA.
AC 066HES;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
OS Hypothetical protein.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Alteschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA Diatchenko L., Marusika K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttenfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC081899; AA81899.1; -;
DR InterPro: IPR011009; Kinase-like.
DR InterPro: IPR00719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE: PS0011; PROTEIN KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
SQ
SEQUENCE 630 AA; 69952 MW; 1626427ABCD5F66E CRC64;
Query Match 93.8%; Score 3087.5; DB 2; Length 630;
Best Local Similarity 94.3%; Pred. No. 3.8e-169;
Matches 595; Conservative 13; Mismatches 22; Indels 1; Gaps 1;
QY 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKHHKHNIRHYEFL 60
DB 1 MESVALLQRPSPQASALASASARPLADGLIKSPKLMKQAVKHHKHNIRHYEFL 60
QY 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 120
DB 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 120
QY 121 EVFENSSKTIYVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIYHRLD 180
DB 121 EVFENSSKTIYVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIYHRLD 180
QY 181 KLENILLDANGNIKADFGLSNLHYHKGFLQTFCCSPLYASPEIYNGKPYGVEVDSWL 240
DB 181 KLENILLDANGNIKADFGLSNLHYHKGFLQTFCCSPLYASPEIYNGKPYGVEVDSWL 240
QY 241 GVLLYTLVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLLMVPTRRATLED 300
DB 241 GVLLYTLVHGTMPFDGQDHKTIVKQISNGAYNEPPKPSDACGLIRWLLMVPTRRATLED 300
QY 301 VASHHVMWVNGYTTGTVGEOBALREGHPSGDFPRASMDWLRSSRPILLENGAYCSPFKQ 360
DB 301 VASHHVMWVNGYTTGTVGEOBALREGHPSGDFPRASMDWLRSSRPILLENGAYCSPFKQ 360
QY 361 HYPGSGSTVPGLEHOSLKKSRKENDMAQNLQGDPAEDTSSRPGKSLKLPKILKKSS 420
DB 361 HYPGSGSTVPGLEHOSLKKSRKENDMAQNLQGDPAEDTSSRPGKSLKLPKILKKSS 420
QY 421 TSSGVEQEDPOLRPPVPTPGQPVPAVSLPRKGIILKKSQRSGYSSPESGELLD 480
DB 421 TSSGVEQEDPOLRPPVPTPGQPVPAVSLPRKGIILKKSQRSGYSSPESGELLD 480
QY 481 ASDVPFSPGDPVEQKSPQASGLILHRKGIILKNGKSRRLBETTPSTGSLDASHPA 540
DB 481 ASDVPFSPGDPVEQKSPQASGLILHRKGIILKNGKSRRLBETTPSTGSLDASHPA 540
QY 541 ARSPRPSGAVSEDSILSSSPDQLPERLPETPLRGCVSVNLAGLQPPSGSLKRWQ 600
DB 541 ARSPRPSGAVSEDSILSSSPDQLPERLPETPLRGCVSVNLAGLQPPSGSLKRWQ 600
QY 601 ESTGDSGSLTDCQEVTAAYRQALGICSKLS 631
DB 601 ESTGDSGSLTDCQEVTAAYRQALGICSKLS 630

RESULT 6
O9H093 PRELIMINARY; PRT; 628 AA.
AC O9H093;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp434J037 (Hypothetical protein SNARK)
DE (Hypothetical protein DKFZp434J037).
GN Name=DKFZp434J037; Synonym=SNARK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RG The German CDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobio G., Han M., Wiemann S.,
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow R.H., Schaefter C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skaleka U., Smalish D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayaishi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AL136891; CAB66825.1; -;
DR EMBL; BC017306; AAH17306.1; -;
DR EMBL; AK074830; BAC11234.1; -;
DR HSSP; P31751; 1GZK.
DR GO; GO:0005524; F.ATP binding; IEA.
DR GO; GO:0004674; P.protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F.transferrase activity; IEA.
DR GO; GO:0004680; P.protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00069; Kinase; I.

DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc.1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 628 AA; 69611 MW; F76F8B1BF94F4C87 CRC64;
Query Match 85.8%; Score 2824.5; DB 2; Length 628;
Best Local Similarity 85.7%; Pred. No. 4.8e-154;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLQPPSQAPSASALASESARPLAAGLISPPPLKKAQVAKGHHKHNLRHYEFL 60
DB 1 MESLVFARRSGPTPS----AAELARPLAAGLISPPPLKKAQVAKGHHKHNLRHYEFL 56
QY 61 ETLGKGTGKVKKARSSGRVLVAKIRKDKIODEDLHIREIRIEMSLNPHIAT 120
DB 57 ETLGKGTGKVKKARSSGRVLVAKIRKDKIODEDLHIREIRIEMSLNPHIAT 116
QY 121 EYFENSCKIVIVMEVSRGDLVDYISERPRLSERDARHFRQIVSALYCHONGIVHRDL 180
DB 117 EYFENSCKIVIVMEVSRGDLVDYISERPRLSERDARHFRQIVSALYCHONGIVHRDL 176
QY 181 KLENILLDANGNKKIADFGLSNLYHKGKFLQTCGSPLYASPEITVNGKPYGPEVDWSL 240
DB 177 KENILLDANGNKKIADFGLSNLYHKGKFLQTCGSPLYASPEITVNGKPYGPEVDWSL 236
QY 241 GVLTYLVHGTVPFDGDHKTIVKQISNGAYREPPKPSACGLIRWLWNVPRRATLE 300
DB 237 GVLTYLVHGTVPFDGDHKTIVKQISNGAYREPPKPSACGLIRWLWNVPRRATLE 296
QY 301 VASHMVMNNGYTTGVGEQALREGGHPGSDFGASADWLRRSSRPLLENGAKVCSFFKQ 360
DB 297 VASHMVMNNGYATRVGEQALREGGHPGSDASASADWLRRSSRPLLENGAKVCSFFKQ 356
QY 361 HYPGGSGTVPGLERQSLKRSKRENDMAONLQDPADETSSRRPKSLKPKILKKKS 420
DB 357 HAPGGSGTTPGLERQSLKRSKRENDMAONLQDPADETSSRRPKSLKPKILKKKS 416
QY 421 TSGGEVQEDPQELRPVDPVPGQVPAVSLPPKGIKKRSORBSGYSSPEPESGELD 480
DB 417 ASHGVQEDPPELSPVPAVPGQVPAVSLPPKGIKKRSORBSGYSSPEPESGELD 473
QY 481 ASDVFPVSGDPEVQKSPQASGLLHKRGIKLNGKFSRTALEGTPSTFGSLDQLASSHPA 540
DB 474 AGDVFPVSGDPEVQKSPQASGLLHKRGIKLNGKFSRTALEGTPSTFGSLDQLASSHPA 533
QY 541 ARPSRPSGAVSEDSIISSSEFPQLDPERLPETPLRGCVSVNLRGLQPPSEGG---LK 596
DB 534 ARASRPSGAVSEDSIISSSEFPQLDPERLPETPLRGCVSVNLRGLQPPSEGG---LK 593
QY 597 RHWQSELDGSCSLTDDCOEVTAAVROALGICRKL 631
DB 594 RWRQDPLGDSCLTDDCOEVTAAVROALGICRKL 628
RESULT 7
O68E04 PRELIMINARY; PRT; 438 AA.
AC O68E04;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp66F01113 (Fragment).
GN Name=DKFZp66F01113;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Fetal kidney;
 RG The German cDNA Consortium;
 RA Othenmaier B., Obermaier B., Deutschenbaur S., Schaap A.,
 RA Mewes H.W., Weill B., Amd C., Oesanger A., Fobio G., Han M., Wiemann S.,
 RA Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CH742209; CH18066.1;
 DR InterPro: IPR011009; Kinase_like.
 DR InterPro: IPR000719; Prot Kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR ProDom: PD000001; Prot Kinase; 1.
 DR SMART: SM00220; S_TKc_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Hypothetical protein.
 KW NON TER
 FT
 SQ SEQUENCE 438 AA; 47528 MW; 6C5E0D89D4D5C4 CRC64;
 Query Match 59.2%; Score 1949.5; DB 2; Length 438;
 Best Local Similarity 84.6%; Pred. No. 5.3e-104;
 Matches 373; Conservative 17; Mismatches 44; Indels 7; Gaps 2;
 QY 195 IADFGSLNLYHKGKFLQTFGSPLYASPEIYNGKPYVGEVDSWLSGLVLYLVHGTMPP 254
 DB 1 IADFGSLNLYHKGKFLQTFGSPLYASPEIYNGKPYTGEVDSWLSGLVLYLVHGTMPP 60
 QY 255 DGDHNTLVKQISNGAYRBPPEKSDACGLIRWLLMVNPTRRATLEVDVASHWVVMGYTTG 314
 DB 61 DGDHNTLVKQISNGAYRBPPEKSDACGLIRWLLMVNPTRRATLEVDVASHWVVMGYATR 120
 QY 315 VEGOEALRGHSGDPPGASADWLRSSRPLLENGAYCSPFKOHVPGGSGTTPGLER 374
 DB 121 VEGOEALRGHSGDPPGASADWLRSSRPLLENGAYCSPFKOHVPGGSGTTPGLER 180
 QY 375 QHSLLKSRKRNMAOMLQGDPAEDTSRPGKSLKLPKGIKKKSSSTSGEVEDEQELR 434
 DB 181 QHSLLKSRKRNMAOMLQGDPAEDTSRPGKSLKLPKGIKKKSSSTSGEVEDEPELS 240
 QY 435 PVPDTFGQVPAVSLPRKGIKKSRQSRSGYSSPEPSGGLDASIVFVSGDVEQK 494
 DB 241 PVPDTFGQVPAVSLPRKGIKKSRQSRSGYSSPEPSGGLDASIVFVSGDVEQK 297
 QY 495 SPQASGLILHRKGIKLNKGFSTALEGTTPTFGSLDQLASSHPARPSPGAVSEDS 554
 DB 298 SPQASGLILHRKGIKLNKGFSTALEGTTPTFGSLDQLASSHPARPSPGAVSEDS 357
 QY 555 IISSEFDDLDLPERLPETPLRGCVVDNLRGLEQPSRS---LKRWQESLGDSCFSL 610
 DB 358 IISSEFDDLDLPERLPETPLRGCVVDNLRGLEQPSRSCLRRMQDPLGDSCFSL 417
 QY 611 TDQCEVTAAYRQALGICSKLS 631
 DB 418 TDQCEVTAAYRQALGICSKLT 438
 RESULT 8
 ARKS HUMAN STANDARD; PRT; 661 AA.
 ID ARKS HUMAN STANDARD; PRT; 661 AA.
 AC 060285;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE AMPK-related protein kinase 5 (EC 2.7.1.37).
 GN Name=ARKS; Synonyms=KIAA0537;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagae T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.

RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 RN [2]
 RP CHARACTERIZATION, AND MUTAGENESIS OF SER-600.
 RX MEDLINE=22393479; PubMed=12409306; DOI=10.1074/jbc.M206025200;
 RA Suzuki A., Kusaka G.-I., Kishimoto A., Lu Y., Ogura T., Lavie M.F.,
 RA Esnault H.;
 RT "Identification of a novel protein kinase mediating Akt survival
 RT signaling to the ATM protein."
 RL J. Biol. Chem. 278:48-53(2003).
 CC -1- FUNCTION: Involved in tolerance to glucose starvation.
 CC Phosphorylates ATM.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- ENZYME REGULATION: Activated by PKB/AKT1 during glucose
 CC starvation.
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in heart and brain,
 CC and at lower levels in skeletal muscle, kidney, ovary, placenta,
 CC lung and liver.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB011109; BAA25463.1; -.
 DR HSSP: P31751; IGZK.
 DR MIM: 608130; -.
 DR InterPro: IPR011009; Kinase like.
 DR InterPro: IPR000719; Prot Kinase.
 DR InterPro: IPR008271; Ser Thr kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot Kinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Phosphorylation; Polymorphism;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 55 306
 FT NP BIND 61 69
 FT ACT SITE 178 178
 FT BINDING 84 84
 FT MOD RES 600 600
 FT VARIANT 543 543
 FT FT
 FT MUTAGEN 600 600
 SQ SEQUENCE 661 AA; 74304 MW; 806F37D52CA4718F CRC64;
 Query Match 51.6%; Score 1699.5; DB 1; Length 661;
 Best Local Similarity 55.5%; Pred. No. 2.1e-89;
 Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;
 QY 14 PSASALASASAPRLDGLIKSPKPLMKQAVYRHHKHLRHYELETLTGKCTYKVK 73
 DB 13 PDLGIGAPSPBRAVAGATPALP-RKPGVKKHHKHLRHYELETLTGKCTYKVK 71
 QY 74 ARBS-SGRVLAISKIKDKIDQODLHTRREIEMSLNPHIILHEVFENSKIVY 132
 DB 72 ATERPSGRVVALISKIKDKIDQODVHTRREIEMSLNPHIILHEVFENKDKIVY 131
 QY 133 MEYASRGDLYDYISERPRLSERDARFFRQIVSALHYCHQNGIVHDLKLENIILDANG 192
 DB 132 MEYASRGDLYDYISERPRLSERDARFFRQIVSALHYCHQNGIVHDLKLENIILDANG 191
 QY 193 IKIADFGSLNLYHKGKFLQTFGSPLYASPEIYNGKPYVGEVDSWLSGLVLYLVHGTM 252
 DB 192 IKIADFGSLNLYHKGKFLQTFGSPLYASPEIYNGKPYVGEVDSWLSGLVLYLVHGTM 251
 QY 253 PFDGQDHNTLVKQISNGAYRBPPEKSDACGLIRWLLMVNPTRRATLEVDVASHWVVMGYT 312


```
DB 252 PFQGFPHKXLIIRISSGEYKEPQPSARGLIRMLMNVNDRKATIEDIANHMMVWNGYK 311
QY 313 TGVGEQDALREGGHPEGSGDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGGSIVPGL 372
DB 312 SSYVDCDADLHDSSEPL-----LARIIDWHHRSTGLQADTEAKMGKLAK---PTTSEVW--L 362
QY 373 EROHSLKKRKENDMAQNLQGDPAEDTSSRPGKSLKLPKGLIKKKS-----STSGEVQ 427
DB 363 ERORSLKKSKKENDFPQSGDAPVES-----PSKLSSRRPGKILKKKSNSHRHSHTGFI 418
QY 428 -----EDPQEL-----RPVDTPGQPPAVS--LLPRKGLIKKKSRORESEY 466
DB 419 GVVGPALPSTFKMEQDLCRTGVLLPSSPEABVPKLSFKPSATMPKKGLIKKQORESEY 478
QY 467 YSSPPESEGEILLDASDVFSVSDPVEQKSP-----QASGLLHRKGLIKLNGKFSRTAL 520
DB 479 YSSPPESESEILLDSNDVM--GSSIPSPSPDPARVYSHSLSCRKGLIKHSGSKYSAGTM 526
QY 521 EGTTPS-----TFGSLDQ-----LASSHPAARPSRPGSAVSEDSILSSSEFPQDLPER 569
DB 537 DPALVSPKEMPTLESISEPGVAPAGLSRSY-----SRRSVISDSDSVLSSDPSFLLDQEN 591
QY 570 LP-ETPLRGCVSVDN-----LRGLRQPSSEGLKRWQBSLGDSCFS-LTDCQEVTAAYR 621
DB 592 RPAPQRIRSCVSAENFLQIDFEGLONRPRPYLKRRLADSSFSLLTMDMDVTQVTK 651
QY 622 QALGICSKLS 631
DB 652 QALGICSKLN 661

RESULT 9
Q61KS PRELIMINARY; PRT; 658 AA.
AC 0641K5;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE RIKEN CDNA B230104P22.
GN Name=B230104P22R1K;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strusberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Pange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gitman J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
```

```
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC082328, AA082328.1, -.
SQ SEQUENCE 658 AA; 73661 MW; E7176f3b75b899f CRC64;
Query Match 51.5%; Score 1695.5; DB 2; Length 658;
Best local similarity 56.0%; Pred. No. 3.5e-89;
Matches 371; Conservative 77; Mismatches 136; Indels 79; Gaps 20;
QY 18 ALASESARPLADGLIKSPKLMKQAVKHHHKNLRHYEFLETIGKTYGKVKARAS 77
DB 26 AVAGATAAPV-----EPR---KPHGVKHHHKNLRHYEFLETIGKTYGKVKARATER 76
QY 78 -GGRVLAISIKDKDKDEODLHTRRELEIMSSLNHPHITAIHEFEKSSKVIYMEVA 136
DB 77 FSGRVVAIKSIRKDKIDELDMWHIRRELEIMSSLNHPHISIYEFENKDKVIIMEVA 136
QY 137 SRGDLVDYISERPRISERDARHPFROIYALHYCHONGVHVDLKIENITLDANGIKYA 196
DB 137 SKGELYDIYSERRRISERETRHFROIYSAVHCNKGVAHDKLENTILDDNCIKYA 196
QY 197 DFGLSNLYHKGRFLQTFCCGSPLYASPEIVNGKRYVGEVDSWSLGLTYLVHGTMPFDG 256
DB 197 DFGLSNLYQKDKFLQTFCCGSPLYASPEIVNGKRYVGEVDSWSLGLTYLVHGTMPFDG 256
QY 257 QDHKTLVQISNGATREPPKESDAGLIRMLMNVNPTRAATLEDVASHWVWVWNGYITGVC 316
DB 257 FDHKNLIRQISSGEYKEPQPSARGLIRMLMNVNDRKATIEDIANHMMVWNGYKSYVC 316
QY 317 EOBALREGGHPEGSGDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGGSIVPGLERQ 376
DB 317 DCDALPDSSEPL-----LARIIDWHHRSTGLQADTEAKMGKLAK---PQASEVW--LERQ 367
QY 377 SLKSKRKENDMAQNLQGDPAEDTSSRPGKSLKLPKGLIKKKS-----STSGEVQ---- 427
DB 368 SLKSKRKENDFPQSGDAPVES-----PSKLSSRRPGKILKKKSNSHRHSHTGFI 423
QY 428 ---EDPQEL-----PVPTPGQVPA--VSL-----LPRKGLIKKKSRORESEY 470
DB 424 PALPSPFKMEQDLCRTAIPLPSSPEADWSGKLSIQSATMPKGLIKKQORESEYSSP 483
QY 471 EPSESGELLDASDVFSVSG-----DPVEQKSPQASGLLHRKGLIKLNGKFSRTAL 526
DB 484 ERSSESEILLDSNDVVISGGLSPPDPARGTSHSLSCRKGLIKKSRSD---GGTDP 540
QY 527 -TFGSLDQASHPAAR-----SRPGSAVSEDSILSSSEFPQDLPERLP-ETPLR 576
DB 541 LRPEMPTLESISPQVPSDGISRSYSRPSIISDVSLSFLLDQENRPARQR 600
QY 577 GCVSVDNLKGL-----QPSSEGLKRWQBSLGDSCFS-LTDCQEVTAAYRQALGICS 628
DB 601 SCVSAENFLQIDFETPHNRPPOYLKR-----LADSSFSLLTMDMDVTQVYKALGICS 655
QY 629 KLS 631
DB 656 KLN 658

RESULT 10
Q61D6 PRELIMINARY; PRT; 575 AA.
AC 061D6;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE MKIAA0537 protein (Fragment).
GN Name=mkIAA0537;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hirooka S.,
```

RA Saga Y., Hagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIA Gene:
RT III. The Complete Nucleotide Sequences of 500 Mouse KIA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.",
RL DNA Res. 10:167-180(2003).
RN (2)
RN SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Nagase T., Ohara O., Koga H.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: AB182364; EMBL:23895.1; -
DR GO: GO:000524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein-tyrosine/threonine kinase activity; IEA.
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase-like.
DR InterPro: IPR00719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00219; TKC; 1.
DR SMART: SM00219; TKC; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER
SQ SEQUENCE 575 AA; 64949 MW; 3CA0796FB4199D2B CRC64;
Query Match 46.0%; Score 1515.5; DB 2; Length 575;
Best Local Similarity 55.7%; Pred. No. 6.4e-79;
Matches 332; Conservative 70; Mismatches 125; Indels 69; Gaps 17;
QY 84 IKSIRKDKTKDEDLHIREIREIMSLNPHIATHEVEFENSKIVIVMEYASRGDLYD 143
DB 1 IKSIRKDKTKDELDVWHIREIREIMSLNPHIATHEVEFENSKIVIVMEYASRGDLYD 60
QY 144 YISERPRLSERDARHFRQIVSAMATYCHQNGIVYHRDLKLNILLDANGIKINDPGLSNT 203
DB 61 YISERPRLSERDARHFRQIVSAMATYCHQNGIVYHRDLKLNILLDANGIKINDPGLSNT 120
QY 204 YHKGKFLQTFCCGSPLYASPEIVNGKPYGPEVDSWISGLVLLYLVIGTMFPDQDHTKYV 263
DB 121 YQDKKFLQTFCCGSPLYASPEIVNGKPYGPEVDSWISGLVLLYLVIGTMFPDQDHTKYV 180
QY 264 KOISNGAYREPPKPSDACGLIRWLLMVPTRATLEEDVASHMVWNGYTTGVGEQALNE 323
DB 181 ROISSEVEYREPTQPSDARGIRKMLMVPTRATLEEDIANHVMWNGYKSSVCDALPD 240
QY 324 GGHPSGDFGRASMAWMLRRSSRPDLLENKAVCSFFKQHVPGGISTVPGLERQHSLSKSRK 383
DB 241 SESPL-----LARIIDWHHRSTGLQAEAEAKMKGLAK---FGASEV---LERQHSLSKSKK 291
QY 384 ENDMAONLGDPRLEEDTSSRPKSSSLKPKGILLKKKS-----STSSSEVQ-----BDPQ 431
DB 292 ENDFPQSGDVSPE-----PSKSSSRPKGILLKRRSSSEHRSHTGTGIVPALPSPF 347
QY 432 ELR-----PYDPTFGQVPVA-VSL-----LPRKGIKKSRQSRGGYSSPEPSSEGR 477
DB 348 KHEQDLCRTAIPFPSSPEADMSGKLSLQSAATMPKGIKKTORRESSGYSSPEPSSESE 407
QY 478 LLDASVFWVG---DPVEQKSPQASGLILHAKGILKLNKFSRTALEGTPS-FTGSLD 532
DB 408 LLDNSNVVVISGGLSSPPPPARCTSHSLSCRKGIKKHGSRYSD---GGTDPALTPPEMP 464
QY 533 QLASHPAARP-----SRPSGAVSEDSIILSESPPDLDLPERLP-ETPLAGCVSYND 583
DB 465 TLESLSPPGVSPGICISRSYISRPSSIIISDSVLSDSFDLDELQENRPARPRISCVSAEN 524
QY 584 LRGLE-----QPSBGLKRWMOBSLGDSCFS-LTDCQEVTAAYRQALGICSKLS 631

DB 525 FQLQDFETPHNRPPOYLKR-----LADSSFLITMDMDVQVYKKALETCSKLN 575
RESULT 11
QY 08CGE1 PRELIMINARY; PRT; 461 AA.
AC 08CGE1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B230104P22Rik protein (Fragment).
GN Name=B230104P22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=salivary gland;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaeetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kerteman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=salivary gland;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC040467; AAH0467.1; -
DR MGD: MGI:1925226; B230104P22Rik.
DR GO: GO:000524; F:ATP binding; IEA.
DR GO: GO:0004672; F:protein kinase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase-like.
DR InterPro: IPR00719; Prot_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
FT NON_TER
SQ SEQUENCE 461 AA; 51274 MW; 0522EC4ECF3713BB CRC64;
Query Match 30.6%; Score 1008.5; DB 2; Length 461;
Best Local Similarity 49.0%; Pred. No. 6.3e-50;
Matches 236; Conservative 59; Mismatches 118; Indels 69; Gaps 17;
QY 198 FGLSNLYHKGKFLQTFCCGSPLYASPEIVNGKPYGPEVDSWISGLVLLYLVIGTMFPDQ 257
DB 1 FGLSNLYQDKFLQTFCCGSPLYASPEIVNGKPYGPEVDSWISGLVLLYLVIGTMFPDQ 60
QY 258 DHKTVKQISNGAYREPPKPSDACGLIRWLLMVPTRATLEEDVASHMVWNGYTTGVGE 317
DB 61 DHGNLIRQISSEVEYREPTQPSDARGIRKMLMVPTRATLEEDIANHVMWNGYKSSVCD 120
QY 318 QEALREGGHPSGDFGRASMAWMLRRSSRPDLLENKAVCSFFKQHVPGGISTVPGLERQHS 377

Db 121 CDALPSESPL-----LARIIDWHRSTGLQABABAKKGLAK---PGASEV--LERQRS 171
QY 378 LKSRKRENDMAONLQGDPAEDTSSRPCKSLKPKIGILKKKS-----STSGEVQ----- 427
Db 172 LKSKSENFPOSGQDSVPES-----PSKLSKRPKGILKRSSEHRSHTGTGIVGS 227
QY 428 --EDPOELR-----PVPDTGQOPVPA-VSL-----LPRKGIILKRSRQREGYSSPE 471
Db 228 ALPSPFKMGODLCRTAIPLPSSPEADMSGSLKQSAITMKKGIILKKTQUREGYSSE 287
QY 472 PEEGGLLDASDVFSVSG---DVEQKSPQASGILLHRKGIILKNGKFSRTALEGTTPS- 526
Db 288 RESESELIDSNVDVIGSLGSSPPDPARGTSHSLGCRKGIILKHSRYSQ---GTDPL 344
QY 527 TFGSLDQLASSHPAAR-----SRPGAVSEDSILSSESTQDLPERLP-ETPLRG 577
Db 345 TRPEMTLESLSPGVPDGSISRSYSPSSIIISDVSLSDSFDLLETQENRPARORIS 404
QY 578 CVSVDTLRGL-----OPPSGLKRMWQESLGSQFS-LTDCOEVTAAVRGALGCSK 629
Db 405 CVSAENFLQLODEPTHNRPQYLKR-----LADSFSLTMDVDTQYKKALETCSK 459
QY 630 LS 631
Db 460 LN 461

RESULT 12
Q65ZH4 PRELIMINARY; PRT; 1551 AA.
ID 065ZH4;
AC 065ZH4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3;
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG MEDLINE=99069613; PubMed=9851916;
RX WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology". The C. elegans Sequencing Consortium.";
RT Science 283:2012-2018(1998).
RN 12
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray U., Le T.T.;
RT "The sequence of C. elegans cosmid B0496.";
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN 13
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN 14
RN SEQUENCE FROM N.A.
AC 065ZH3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein B0496.3;
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 15
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG MEDLINE=99069613; PubMed=9851916;

DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
SQ SEQUENCE 1551 AA; 176395 MW; 1F3BE62632A44C95 CRC64;
Query Match 27.0%; Score 888; DB 2; Length 1551;
Best Local Similarity 37.0%; Pred. No. 2, 6e-42;
Matches 230; Conservative 73; Mismatches 214; Indels 104; Gaps 15;
QY 54 RHRYEFLTGLKGTGKVKARESS-GRVVAIKSTRKDKIKDQODLHIREIETSSIN 112
Db 72 KRFETTKGSGGTGKSLAYDHKFDREVAALIKKSAIESKADLVRIREIRISALN 131
QY 113 HPHIAIHVEFNENSKIVMEYASRGDLYDIISERPRLSERDARFFRQIVSALYCHQ 172
Db 132 HNNIIQIVVFENKKIILVMEYSSGGLYDYVSRGSLPEAARIRFQIISAVLYCHK 191
QY 173 NGIVHDLKLENIILDPANGNIKIADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVG 232
Db 192 HRVAHRDLKLENIILDPQNNNAKIADFGLSNYPADKULLTTFGSGSPLYASPEIINGPYNG 251
QY 233 PEVDSWSLGLVLYLVHGTMPDGDHKTLYVQISNGAVREPPKPSDAGLIRWILMNP 292
Db 252 PEVDCWSLGLIYLTYLVGSMPPDGRDPRMVMQIKRGAFIEPETSTASMLIRNMLRVNP 311
QY 293 TRRATLEDVASHMVMVNGYTTGV-----GEQALREG 324
Db 312 ERRATIFDASHMVLNLENMVPVIOELPENQIIDHTPLTEREETMIVODLADQDVFMF 371
QY 325 GHPSGDFGASAMDLR-----SRPPLLENAGKCSFKQIVPGGGSTV 369
Db 372 GHLSSETRKIKDFIRKKEAEFNDNSPVKPKARKTDELGKISKEQPEEKSAEKL 431
QY 370 PGLERQHS-----LKSRK-ENDMAONLQGDPAEDTSSRPCKSLKPKIGILKKK 418
Db 432 RGKKEKEKPKVVDPNDFLRQIENRIGQKKKQKAEAKTARVETVKL-----EVK 485
QY 419 SSTSGEVOEDPOELRPVDTGQOPVPAVSLPRKGIILKRSRQREGYSSPESSGEL 478
Db 486 KEKSPQGPEDPKTAR-----GTSKPADSRAP-SFVPYKDPET---SEERPRTRH 534
QY 479 LDASDVFSGDPVEQKSPQASGILLHRKGIILKLN-GKFSRTALEGTTSTFGSLDQLAS 536
Db 535 LTPASAVRIETDILNMLMNV--LEQMEKGPVMIITARIKAPLYDTRMVKELLESIA 592
QY 537 SHPARPSRPGAVSEDSILSSESPDL---DLPERLPETPLRGVSDNLRGLQEPS 592
Db 593 AQPEVQKQTSKVEEQFTSRONTLTRKKEDPLEEIEVP-----SPS 637
QY 593 EGLK-RWV-----QESLG 604
Db 638 RKKKERPMHSVAVCKMNSVG 658
RESULT 13
Q65ZH3 PRELIMINARY; PRT; 1592 AA.
ID 065ZH3;
AC 065ZH3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3;
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 16
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG MEDLINE=99069613; PubMed=9851916;

RG WormBase Consortium;
RT "Genome sequence of the nematode *C. elegans*: a platform for
RT investigating biology. The *C. elegans* Sequencing Consortium.";
RL Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Br1etol N2;
RT Murray J., Le T.T.;
RT "The sequence of *C. elegans* cosmid B0496.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Br1etol N2;
RA Watson R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Br1etol N2;
RG Wormbase Consortium;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58749; AAU05598.1; -.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KM Hypothetical protein.
SQ SEQUENCE 1592 AA; 180962 MW; E28D6FF934C3F76 CCK64;
Query Match 27.0%; Score 888; DB 2; Length 1592;
Best Local Similarity 37.0%; Pred. No. 2,7e-42;
Matches 220; Conservative 73; Mismatches 214; Indels 104; Gaps 15;
QY 54 RHRYEFLTLGKGTGYKVKARESS-GRVAISIRKDKTKBODLHTRREIEMSSLN 112
DB 72 KHRFEFTKLGSGTYKVSILAYDHKFDREVAVLLIKKSAIESKADIVRIREIRIMSALN 131
QY 113 HPHIILHEVEENSRTIVMEYASRGDLYDYISERPRLSERAPRHFQIVASLHYCHQ 172
DB 132 HPHIILQIYEFPEKKDKITLMEYSSGGLDYDVSRCGSLPEARARRIFRQITSAVLYCHK 191
QY 173 NGIYHDLKLENTILDANGNIKIADFGLSNLYHKGFLOTFCGSPYASPEIYNGKPYVG 232
DB 192 HRYAHDLKLENTILDONNNAKIALDFGLSNLYFPADKXLLTFCGSPYASPEIINGTPYKG 251
QY 233 PEYDWSLGLVLLIYHGTWPFPGDQHKTLVKQISNGAVREPPKPSDACGLIRMLMVP 292
DB 252 PEYDCSLGILTYLVYGSWPFGRDPFNRMVROIKRGAYEPPPTSTASWLIHNMRLVNP 311
QY 293 TTRATLEDAVASHMMVMWGGTTGV-----GGEALREG 324
DB 312 ERRATIFDIASHMWMLLENMPTIQLPENQIIDHTPLTERBETMIVQDLADQDFMER 371
QY 325 GHPSGDFGRASMDMLRR-----SSRPLENGAKVCSFPGOHPPGGGSTV 369
DB 372 GHLSSTRKRIEPRIRKKEAEFPDNPSPYKPRAKRTDBLTGKISKEGEBKSKAEKSL 431
QY 370 PGLERHNS-----LKKSRK-ENDMAQNTQGPDAEDTSSRPKGSILKLPKGLIKKK 418
DB 432 RGVEKEKEKPKVVDPNDFLERLQIENRLGQNKCKDEAAKTSARVETVKLK-----EVK 485
QY 419 SSTSGSEVGEVDEPGLRVPPTPGQVPVAVSLPRKGLIKKSRREGSYSSPPESGEL 478
DB 486 KEKSPQOQEPDPTAR-----STSKPADSRAP-SFPVVDREPT-----SEPRPRTRPP 534
QY 479 LDASDVFSVGDPEVOKSPQASGLLHRKGLIKLN--GKFSRTALEGTPSTFSLDQLAS 536

DB 535 LRASAVRITDLSNMLMNQV--LEONKEGPVNIIARIKAMPLDYTRPVPKELLSITA 592
QY 537 SHPARPSRPSGAVSDESISSSESPDL-----DLPERLPTPLRGCVSYNDNLGRQPPS 592
DB 593 AGPEPVOKOTSKVVEQOTRSRONTLTRKKKEDPLESEIFBPV-----SPS 637
QY 593 EGKLR-RWV-----QESLIG 604
DB 638 RKKKERPMHSVEVCMKNESVIG 658
RESULT 14
ID Q7KSS0
AC Q7KSS0;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE CG11870-PD.
GN ORFNames=CG11870;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Cham M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R.K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklow G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klatas I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodagat, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhu Q., Zheng L., Ye J.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirkas R., Tabor P.E., Man K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence."
 RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirkas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celisner S.E.,
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective."
 RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk K.S.,
 RA Smith C.D., Tupy J.L., Whitfield E.D., Bayraktaroglu L., Beiman B.P.,
 RA Belcencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review."
 RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG Flybase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AB003686; AAS6535.1; -.
 DR HSSP; Q63450; 1A06.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase_like.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 RN KW
 SQ SEQUENCE 1180 AA; 127743 MW; A88A012664046EC6 CRC64;

Query Match 26.8%; Score 881.5; DB 2; Length 1180;
 Best Local Similarity 32.7%; Pred. No. 4,3e-42;
 Matches 238; Conservative 101; Mismatches 216; Indels 173; Gaps 21;

QY 6 ILQRP-SQAPASALASGARPLADL-----IKSPKLMKKQA-----VXR 46
 DB 2 VLSKPDGTAPNGAAGAEAAAP--TGLDATGNSLHPSGIPQDQIDNMSGIANTGNVYM 59
 QY 47 HHNKNLRRRYEPLTGLGKTYGKAKKA-RESSGRVLAIKSRKDKIKDQDLHTRRI 105
 DB 60 NNHRRKRLRFIDIKKGGTYGKQVLGINKETGQVAILTKIKCKIEAEADLVRRREV 119
 QY 106 RIMSSLNHHPIAIHIEVFENSSKIVVMEYASRGDLVDYDISEPRPSEADAHFFQIYS 165
 :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

DB 120 QIMSSVHHNIIHIEVFENRREKMLVMEFAAGSELYDLSRKVLTEEARIRFQVAT 179
 QY 166 ALHYCHQNGIYVRDLKLENTLIDANGNTKIADFGSLNLYHKGFLQTFGSPLYASPEIV 225
 DB 180 AVYCHKHKICHRDLKLENTLIDENGNAKIAFGSLNVPDDRLQITFGSPLYASPEIV 239
 QY 226 NKPVPYGPVDSMSGLVLYIWHGTMPPDGDKTLVYQISNGVAREPXPDSADGLR 285
 DB 240 EGTPIYGPVDSMSGLVLYIWHGTMPPDGDKTLVYQISNGVAREPXPDSADGLR 299
 QY 286 WLLMNPTRATLEDAVSHMWVN-----WGVTGVGQEAALR-----E 323
 DB 300 DWLTVCPRRBAIEIQCSHMMWNENDNVSCLDAEDLANQTVRDVLLSLPATITTAQ 359
 QY 324 GGHPSGDFGRASMDWLRSSR-----PLENGAVYCSFFKQHPGGS---- 367
 DB 360 LVVPSAEGAAAKAANERVPFRSHVGSIRDMGPPTTEARRLLDM-----VAAAGEAALM 415
 QY 368 -----TVPLEGQHSLSKSRKENDMAQNIQGPADETSRRPKSSL-----KLPR 412
 DB 416 PSPRTITTPAQSPVQTKRLQPTVSTENAAATTAK-KKEKPANSSFVLSKDGAPLLEADP 474
 QY 413 GILKKSSSTSGE-VOEDPQELRPVPD-----TPGQPVPA-V 447
 DB 475 TIIEPGATLMEAFITANIPDEV-PVPSYSQKMDQAVGDCILMGATTPAPAPAPTPV 533
 QY 448 SLIPRKG-----ILKK----- 458
 DB 534 ARQPTRGKLDVAVETPEEKDAATKVIKKFNKHTADLVNAINESAKAAPVSAVAPPF 593
 QY 459 -----SRQPSGYSSPPEBSEGLDADSVFVSGPVPKESQ-ASGLLHKKGIKLNG 513
 DB 594 VRKSLQDSTLKNKNAERRKSRILETAEKFPPEPPVAAAPPEKPKKLSIPGVSVGSPFK 653
 QY 514 KFSR-----TALEGTTPSTFGSLDQLASSHPAR-----PSRPSGAVSSEDSILSSS 560
 DB 654 EBEKATNPAAEGTPEGLRQEQVAAAAAQAELTSTPSPVVAQSLGSDSKS 713
 QY 561 FPDLDLPE 568
 DB 714 VASISIDE 721
 RESULT 15
 QYU19 PRELIMINARY; PRT; 1180 AA.
 AC QYU19;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE LP05937P.
 GN ORFNames=CG11870;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidae; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Broksstein P., Hong L., Agbayan A., Carlson J.,
 RA Change M., Chavez C., Dorsett V., Dresnek D., Fattán D., Frise E.,
 RA George R., Gonzalez M., Guarín H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Pargae V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.,
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BT010033; AAC2502.1; -.
 DR HSSP; P31751; 1MRV.
 DR Flybase; FBgn0037804; CG11870.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

Job time : 167.033 secs

DR GO:0016740; P:transferase activity; IEA.
DR GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR Pfam: PF00069; PKinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1180 AA; 127684 MW; 2CD790A4F3D8C50C CRC64;

Query Match 26.8%; Score 881.5; DB 2; Length 1180;

Best Local Similarity 32.7%; Pred. No. 4,3e-42;

Matches 238; Conservative 101; Mismatches 216; Indels 173; Gaps 21;

QY 6 LQRP-SQAPSAALASASARPLADGL-----IKSPKPLMKQA-----VKR 46
DB 2 VASKPDGTAPNGAGAAAP--TGLDATINGLAPSGIPDQIDNMSGINTGNVKK 59
QY 47 HHKKNLRHRYEFLTYGKTYGKKA-RESSGLVAIKSIRKDKIKDQDLHIRREI 105
DB 60 NNHRKRLRQRFDIKKLGQGYGVQVGLINKETGQVAKTKICKIEADLVIRREV 119
QY 106 EIMSSLNHPHIIAHEVFENSSKIVMEVYASGDLVDYISERPLSERDARHFFROIYS 165
DB 120 QIMSSVHHNPIIHYVFENREKCVLWEPFAAGGELYDYISERKVLTEBARIRFQVAT 179
QY 166 ALHYCHONGIVHRDLKLENIILLDANGNIKIADFGLSNLVHKGFLOTFCSPLYASPEIV 225
DB 180 AVYCHKHKICRDLKLENIILLDEKNAKIDRGLSNVFDQRLGTFGSPLYASPEIV 239
QY 226 NKPVYGPPEVDSMSLGVLLYILVHGTPPPDGDQDKTLVQISNGAYRPPKPSDACGLIR 285
DB 240 EGPYQGPPEVDCSLGLVLLYTLVYSGMPDGSNFKRLVQISQGDYRPPKPSRASTLIR 299
QY 286 WILLMNPTRRATLIEDVASHMWN-----MGTTGVGEQEAR-----E 323
DB 300 DMLTVCRRRASTEQICSHMWNENDNVSCLDLAEDLANQTPVRLDVLSLTPATTADQ 359
QY 324 GHPSGDFGRASADWLRRSSR-----PLENGAKVCSFQKHVPGGS----- 367
DB 360 LVVPSAEGAALAAANERVPKSHSVGSTRDMGPENTEARILDM-----VAAGGAALM 415
QY 368 ----TPGLEHQSLKSKRKENDMAQNTQGDPAEDTSRPGKSL-----KLPK 412
DB 416 PSPTRTITPAQSPVQTKRLQPTVSTENAGTTAK-KKEKPANSSFVLSKDGAPLTEAPP 474
QY 413 GILKKKSSSTSGE-VEEDPOELRPVPD-----TRGQVYPA-V 447
DB 475 TIIEPQATIMEATIANIPEEV-FVPSYSQDMQAVDICDILMGATTAPPAPAPPTPV 533
QY 448 SLPRKG-----ILKK----- 458
DB 534 AROPTRGKLDVAVETPEEKDATKVIKKFVNKHTADLVANINESAKAAAPVSAVAPPP 593
QY 459 ----SRQRESGYVSSPEPESGELLDASDVFNVSQDPEVQKSPQ-ASGLLHRRKGIKLG 513
DB 594 VRKCSLQDSTLTKFNAERKRSRLLETAKEFQPPVAAAPPEKPKKLSIPGVSQSPKK 653
QY 514 KFSR-----TALGTTPTSTGSLDQIASSHPAAK-----PSRPGAVSEDSIISSES 560
DB 654 EPEKATNPAAAGPTPGLRAQGVAAAAAAQAEBETLSTPPSPVVAQSLGSDSKNS 713
QY 561 FDQLDLP 568
DB 714 VASISIDE 721

Search completed: May 11, 2005, 14:24:00

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:13:08 ; Search time 38.5452 Seconds
(without alignments)
1575.104 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293

Sequence: 1 MESVALQRPQSPQASASALA.....DCQEVTAAYRQALGICSKLS 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_79: *
2: pir1: *
3: pir2: *
4: pir3: *
5: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	888	27.0	1558	2 T92953	hypothetical prote
2	694.5	21.1	713	2 S27966	probable serine/ch
3	689	20.9	745	2 G01025	serine/chreonine-s
4	686.5	20.8	774	2 T48609	probable serine/ch
5	669	20.3	1192	2 T18611	-probable serine/ch
6	667	20.3	1246	2 G89287	protein H39E23.1 (
7	664	20.2	1398	2 T13741	hypothetical prote
8	649.5	19.7	512	2 T52633	serine/chreonine-s
9	648.5	19.7	633	1 A26030	serine/chreonine-s
10	648	19.7	798	2 JC7500	qik protein - chic
11	641	19.5	504	2 T10449	probable serine/ch
12	639.5	19.4	512	1 JCI446	serine/chreonine-s
13	631	19.2	511	1 A56009	serine/chreonine-s
14	630.5	19.1	552	1 S51025	[hydroxymethyl]glu
15	627.5	18.9	602	2 S72513	FOG2 protein - yea
16	623.5	18.9	552	1 A53621	[hydroxymethyl]glu
17	622	18.9	562	2 T29858	hypothetical prote
18	621	18.8	472	2 B90100	SNF-related kinase
19	620	18.8	504	2 T07415	probable serine/ch
20	612	18.6	481	2 T49072	protein kinase
21	610	18.5	473	1 S59941	serine/chreonine-s
22	599	18.2	512	2 T07788	probable serine/ch
23	596.5	18.1	887	2 T20941	hypothetical prote
24	594.5	18.1	576	2 T41587	probable carbon ca
25	587.5	17.8	622	1 S44859	serine/chreonine-s
26	583.5	17.7	891	2 A38903	protein kinase 1 -
27	581.5	17.7	891	2 T40503	protein kinase kin
28	579	17.6	510	2 T04145	serine/chreonine p
29	570	17.3	502	1 A41361	serine/chreonine-s

30	569.5	17.3	401	2 B90120	SNF1-related prote
31	564	17.1	513	1 S60303	serine/chreonine-s
32	563	17.1	513	1 S60304	serine/chreonine-s
33	559	17.0	1147	2 S64930	serine/chreonine-s
34	549.5	16.7	441	2 C84667	probable protein k
35	549	16.7	1064	2 S52687	serine/chreonine-s
36	547	16.6	1142	2 S59359	G1N4 protein - yea
37	545	16.6	651	2 S52444	p69Eg3 protein - A
38	543.5	16.5	480	2 A86427	probable serine/th
39	543	16.5	800	2 S29344	protein kinase KIN
40	539.5	16.4	440	2 T14735	probable serine/th
41	537.5	16.3	440	2 T14736	probable serine/th
42	534.5	16.2	435	2 E84707	probable protein k
43	529.5	16.1	726	2 T33998	hypothetical prote
44	527	16.0	746	2 S62365	SNF1-related prote
45	524	15.9	502	2 T02306	probable protein k

ALIGNMENTS

RESULT 1	
T92953	hypothetical protein B0496.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans	
C/Date: 15-Oct-1999	#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T92953	
R/Murray, J.; Lee, T.T.	
Submitted to the EMBL Data Library, May 1996	
A/Description: The sequence of C. elegans cosmid B0496.	
A/Reference number: Z20596	
A/Accession: T92953	
A>Status: preliminary; translated from GB/EMBL/DBJ	
A/Molecule type: DNA	
A/Residues: 1-1558 <MDR>	
A/Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3	
A/Experimental source: strain Bristol N2; clone B0496	
C/Genetics:	
A/Gene: CESP:B0496.3	
A/Map position: 4	
A/Introns: 85/1; 100/3; 141/1; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2;	
Query Match	27.0%; Score 888; DB 2; Length 1558;
Best local similarity	37.0%; Pred. No. 2.1e-28;
Matches	230; Conservative 73; Mismatches 214; Indels 104; Gaps 15;
QY	54 RRRYFLETLGKGTGKVKKARESS-GRIVAISKIKDKIKDQDILHRRREIINSSLN 112
DB	72 KRRFEITTKLGGTGYKSLAVDHKFDREVAVKLKKSAIESKADIVIRREIRIMSA 131
QY	113 HHHTIAIHVEFENSKIIVVMYASRGDIYDYSRPRLESDAHFFQIYSALHYCHQ 172
DB	132 HNNITQIYEVFENKCKITILVMEYSSGGELYDVASCGSLPEAEARIFQIYSAVLVYCK 191
QY	173 NGIVRDLKLENIILLDANGNIKIADFGLSNLHYHKKFLOTFCGSPPLYASPEIVNGKPYVG 232
DB	192 HVAHRDLKLENIILLDQNNAKIADFGLSNVRADKNLITTFGSPPLYASPEIINGPYVG 251
QY	233 PEVDSWSLGVLLIYLIVHGTMPDGDHKTLYVQISNGAVREPPKPSDAGLIRMLLVNP 292
DB	252 PEVDCMSIGIILTYLVYSGMPDGRDFNMVQIKRGAVFEETSTASMLIRNMLRVNP 311
QY	293 TTRATLEVDASHWVNWVWGTTGV-----GEQELRRG 324
DB	312 ERRATIFDLASHWMLNLEENMPVIOELPENQIIDHTPLTEREETMIVODLADQDVFMEP 371
QY	325 GHPSGDFGRASWADLRR-----SSRPLLENAGKYCSFFKQHVPGGGSYV 369
DB	372 GHLSEETRRKIDFPIRRKEAEFPNDNSPVKPKARKDELNGKISKQPEEMKSKAKSL 431
QY	370 PGLERQHS-----LKSKRK-ENDMAQNLQGDPAEDTSRPSKSSLKPKPIGLKK 418
DB	432 RGVEKEKEKPVVDPNDPLERLQIENRLGQNKQKEAAKTSARVETVTLK-----EVK 485

```

QY 419 SSTSGGEVQGEPOELRPVPDTRGQVPVAVSLPRGILKTKRQRESGYSSPEPSESEL 478
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 486 KEKPEEQEEDPEPKTAR-----GTSKPADSRAP-SFVVKORPET---SEPERRTTPE 534
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 479 LDASDVFSQDPVEOKSPQASGLLLHRKGIKLN--GKSRTALEGTPTTFSGIDOLAS 536
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 535 LTAAYRRETDLSNLTNQV--LEGMEKGPVNLINIAIKHPLDYTRPMKELLESTIA 592
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 537 SHPARPSRPGAVSEDSILSSSEFDQ---DLPERLPETPLRGVSVNDMLRGLBOPS 592
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 593 AQPEPVQKQTSKVVEQQTFSRQNTLTTRCKEDPLEEBIEPV-----SPS 637
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 593 EGLK-RMW-----QESLG 604
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 638 RKMKERPMHSAVECMKRESVG 658
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 2
 S27966
 Probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human
 N/Alternate names: protein p78
 C/Species: Homo sapiens (man)
 C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
 C/Accession: S27966
 R/Maheshwari, K.K.; Som, S.; Parsa, I.
 submitted to the EMBL Data Library, January 1992
 A/Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced
 A/Reference number: S27966
 A/Accession: S27966
 A/Molecule type: mRNA
 A/Residues: 1-713 <MAH>
 A/Cross-references: UNIPROT:P27448; EMBL:M80359; NID:g189511; PIDD:AAAS9991.1; PID:g189599
 C/Superfamily: protein kinase homology
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 C/54-307/Domain: protein kinase homology <KIN>
 C/62-70/Region: protein kinase ATP-binding motif

Query Match	21.1*	Score 694.5;	DB 2;	Length 713;
Best Local Similarity	34.5*	Pred. No. 4.9e-21;		
Matches 191;	Conservative 75;	Mismatches 171;	Indels 117;	Gaps 17;
QY	YEPLETTGKGYGVKKAR-	SSGRLVAKIRKDKDEODLHIREITMSLNPH	115	
DB	57	YRLAKTIQKGNFAVKLARHLLTGREVAIKIIDTQL-NPISLQKLFREVIIMKILNPN	114	
QY	116	IIAHEVFENSKKIVIMEVASRGDLYDYSERPLSERDAHFROIVSAIHYCHONGI	175	
DB	115	IVKLFVEIENQKTYLIMEYVSGGKVPDYLAHGRMEKEKRSFRQVSAVQYCHOKRI	174	
QY	176	VHRDLKLENIILDDANGNIKIADFGLSNLVYHKGKTLQTFCCSPPLASPEIYNGKPYGPEV	235	
DB	175	VHRDLKLENIILDDANGNIKIADFGFSNFEVTGGKIDTFCCSPPLAABELFGKKYDGPPEV	234	
QY	236	DSWELGLVLYLVHGMTPFGODHKTIVKQISNGAYEPRPKSPAC-GLIRMLLNWPTK	294	
DB	235	DWSLSGLVLYLVHGMTPFGODHKTILREYLRKRYRIPFYMSIDCENTLRRFLVLPNIK	294	
QY	295	RATLEDAVSHWVMGYTTGVGEQALREGHPSGDFGRASMAAMLRSSRPLLENGAKV	354	
DB	295	RGLTEQIMKORMINAGH-----EDELKPFYEPBELDISDQKRITIM-----V	336	
QY	355	CSFEKQHVPGGASTVPELGRQHSILKSKRKENDMAQNL-----QGDPAEDTSRSPKGS	407	
DB	337	GMGYSQE-----EIGESLSKMKYDEITATYLLLGKSKSEVAPRSSLNINSTQGS-	364	
QY	408	LKLPKGLIKKKSSTSSGGEVQEDPQELRPVPTPQQVPVAVSLFR-----	452	
DB	385	---PHHVVQRSSVSS-----OKORRYSDHAPGIPSVVAVYKRSQTSYADSLKEDG	433	
QY	453	-----GILLKSRQRESGYSSPESPESGEHLDAADVPVSGDVEQSKSQAS	499	
DB	434	ISRSKTSGLAVGGGILPASPML-IGNASINENKADIERKKSSSTV-----PSSWTAS	483	

```

QY      500 GILHHRK-----GILKNGKPRSRALGCTPTSTGSLDQLASH--PAAP 543
Db      484 GGMTRRNTYVVCSEKRTTDDRHSVIO-NGKENSTIPPORT-----VASTHSISSMTP 534
QY      544 SR---PSGAVSEDS 554
Db      535 DRIRPRGTASRST 548

```

```

RESULT 3
G01025
serine/threonine protein kinase - human
C|Species: Homo sapiens (man)
C|Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C|Accession: G01025
R|Navarro, E.
submitted to the EMBL Data Library, April 1996
A|Reference number: H00564
A|Accession: G01025
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1-745 <N>
A|Cross-references: UNIPROT:Q15524; UNIPROT:Q96H83; EMBL:X97630; NID:G1310674
C|Superfamily: protein kinase homology
F18-271/Domain: protein kinase homology <KIN>

```

Query Match	20.9%	Score 689	DB 2	Length 745
Best Local Similarity	33.2%	Pred. No. 8,4e-21		
Matches 189	Conservative 92	Mismatches 176	Indels 112	Gaps 20
QY	57	YEFLETLCKGYGYKKKARE--SSGLVAIAIKSIRKDKIDKEODLHRIREIEMSSINHPH	115	
DB	20	YRLKTIKGNPAKFKLARHLITGKEVAVKIDKQL--NSSSLQKLFRVRLIMKVLNHN	78	
QY	116	IIAHEVENSCKIYIVNEIYASRGDLYDTSRRPRLSEBDANHFRQIYVSALHYCHQNGI	175	
DB	79	IVKLFEEVLETEKTLVLVMEIYASGGEVFDYLVHGMKKEKAPKFOIYSAVOYCHQKFI	138	
QY	176	VHRDLKLENIILLDANGNIKIDPFGLSNLYHKKFLQTFGSGPLVYASPELVNGKPYVGPEY	235	
DB	139	VHRDLKLENIILLDADNMIKIADFGSGNNEFTFGNKLDITFGSGPPYAPFLFGKKYDGPY	198	
QY	236	DSWSLGLVLYLLVHGTMPEFDQDHNTLYVKQISNGAYREBPKEPSDAC--GIIRMLMVNPT	294	
DB	199	DWSLGIYVILTYLVSGLSPFDDQNLKELREBYLRGKRIIPFYWSTDCENLLKFLILNPSK	258	
QY	295	RATLEDVASHHVMVMGYTTGVGEQALREGGHPSPDPGRASADMLRRESRPLENGAY	354	
DB	259	RGTLEQIKDKRMWNVGH---EDDELKPYVEPLPDYKDP-----RTLELWVSG---	303	
QY	355	CSFPGQHPVGGSTVPLGEROHSLEKKSREKDMA-----ONLQGD-----PAED	398	
DB	304	--YRREI-----QDSLVGQGYNEMVATYLLGYKSSLEEGTITLKKRPSAD	349	
QY	399	TSSRPKGSLSLKLPRGI-----LKKSSSTSGE---VOEDPOE	432	
DB	350	LTNSSAOPSPSHKVRASVANSPKORRPSDOAGPAIPTSNYSKTKTOSNNAMENKRPEDRES	409	
QY	433	LRPVPDT---GQPPAVASLLPRKILKKSRQREGGYSSPPRPSGGLLDASDVVS	489	
DB	410	GRKASSTAKVVASPLP-----GLEKK-----TTPPSTNSVLSSTIN--	447	
QY	490	PVEKSPQASGLLHRKGIKL--NGKFSRTALEGTTPTFGSLDOLASSHPAARPSR	546	
DB	448	-RSHNSP-----LLERASLGQASIQNGKDSIT--MGSRASSTASAGVGAARPRQKSM	500	
QY	547	SGAV---SEDSILSESFDQLDPRLE	572	
DB	501	SGSVHPNKASGLPTESNCEVPRSTAQ	529	

RESULT 4

148609
Probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #ext_change 16-Aug-2004
C:Accession: 148609; S31333
R:Inglis, J.D.; Lee, M.; Hill, R.E.
Mamm. Genome 4, 401-403, 1993
A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
A:Reference number: 148609; MUID:93364122; PMID:8358177
A:Accession: 148609
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-774 <ING1>
A:Cross-references: UNIPROT:Q05512; EMBL:X70764; NID:957919; PIDN:CA5040.1; PID:957920
R:Inglis, J.D.; Lee, M.; Hill, R.E.
submitted to the EMBL Data Library, January 1993
A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome
A:Reference number: S31333
A:Accession: S31333
A:Molecule type: mRNA
A:Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTSGTCYAGNAHQATRT', 731-774 <ING2>
A:Cross-references: EMBL:X70764
C:Genetics:
A:Gene: emk
C:Superfamily: protein kinase homology
C:Keywords: ATP, phosphotransferase, serine/threonine-specific protein kinase
F:51-304/Domain: protein kinase homology <KIN>
F:59-67/Region: protein kinase ATP-binding motif

Query Match 20.8%; Score 686.5; DB 2; Length 774;
Best Local Similarity 32.6%; Pred. No. 1.1e-20;
Matches 199; Conservative 90; Mismatches 218; Indels 103; Gaps 20;

QY 6 ILQSPQAPASALASESARPLADGILKSPKPLMKQAVRHHKKNLHNR-----YEEL 60
DB 11 LNERTEQPTLGLHDS-----KPSKSNMLRGRRSATSDQPHIGYRI 56
QY 61 ETLGKTYGKVKKARE-SSGRVLAIKSRKDKIKDEODLLHIREIEMSLNHPHII 119
DB 57 KTIQGNFPAKVLAAHILILGKEVAVKIIDKTQL-NSSLOKLFREVRIMKVLNHPHIVL 115
QY 120 HEVFNSSKIVIMEYASRGDLYVIERPRLSERDARHPFQIYSALHYCHONGIVHD 179
DB 116 FEVIERTEKLYLIMEYASGGEVFDYLVAGRMKEKRAKRFQIVLVQYCHQKFIYHD 175
QY 180 LKLENIILDANGNIKADGLSNLHKGFLOTFCSSPLVYASPEIYNGRPYGPEDSMS 239
DB 176 LKAENILLADADNMKIKADFGFNSLFFGKLDLTFCCSPYAAPBELFQGGKIDGPEVDVMS 235
QY 240 LGLVLYLVHGTMPPDGDHKTLYKQISNGAYREPPKPSDAC-GLIRMLMVNPTBATL 298
DB 236 LQVILYLVVSGSLPFDGQNLKELREYVLRGKTRIPFYMTDGCNLKFLIILNPSRGTL 295
QY 299 EDVASHMVMVWGTYTGVGEQALREGHPSGDFGRASMDWLRSSRPULLENGAKVCSFF 358
DB 296 EQIMDRMWN-----VGHED-----DELKPYEPLLTGPR----- 325
QY 359 KQHVGGGSGTVFGL--ERQHSIKSRKKNEMAQNL-----QGPADETSSRPKSSSLQ 410
DB 336 RDRVVG---VNGLHTEBIDSLVGRVYEVNATYLLIGYKSEPPGDTITTLKPRSSADL 381
QY 411 PGILKKKSSSTSGVEQDPOELRPVDPGQVPVAVSLPRKGLIKSRQESGVYSSP 470
DB 382 TMSASPSBHKVQSVSNPKQRR---SSDAVPI---PTSNYSKTKQTSNNAENKRP 434
QY 471 EPESEGLLDASDVFGDP-VEQK---SPQASGL-----LIHRKGLIKL-- 511
DB 435 EBEETGRKASTAKVPASPLPGLDRKXTTAPSTNSVLTSTNRSRSPGLIDRASLQASI 494
QY 512 -NGKTSKTLLEGTPBTFSSLDGLASHP-----AARPRPGAVSEDSILSSSEF 561
DB 495 QNGKDSLT-MPGSRASTASASAVSARPRQHKQMSASVHPKASGLPPTES--NCEVP 551

QY 562 DQDLPERP 571
DB 552 RPSRAPORVP 561

RESULT 5
T18611
Probable serine/threonine-specific protein kinase (EC 2.7.1.-), long splice form - Caen.
N:Contains: probable serine/threonine kinase, short splice form
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C:Accession: T18611; T18610; T23144; T23143
R:McMurray, A.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z18997
A:Accession: T18611
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W113>
A:Cross-references: UNIPROT:Q9TW45; EMBL:Z81027; PIDN:CA54179.1; GSPDB:GN00023; CESP:H3
A:Experimental source: clone AH10
A:Accession: T18610
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W12>
A:Cross-references: EMBL:Z81027; PIDN:CA54178.1; GSPDB:GN00023; CESP:H39E23.1b
A:Experimental source: clone AH10
R:McMurray, A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z19696
A:Accession: T23144
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W113>
A:Cross-references: EMBL:Z96102; PIDN:CA54263.1; GSPDB:GN00023; CESP:H39E23.1a
A:Experimental source: clone H39E23
A:Accession: T23143
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W14>
A:Cross-references: EMBL:Z96102; PIDN:CA54262.1; GSPDB:GN00023; CESP:H39E23.1b
A:Experimental source: clone H39E23
C:Genetics:
A:Gene: CESP:H39E23.1a; CESP:H39E23.1b
A:Map position: 5
A:Insertions: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 98/
C:Keywords: alternative splicing; ATP, phosphotransferase, serine/threonine-specific pr
F:1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #
F:1-487, 536-1192/Product: probable serine/threonine-specific protein kinase, short spli

Query Match 20.3%; Score 669; DB 2; Length 1192;
Best Local Similarity 32.4%; Pred. No. 8.2e-20;
Matches 187; Conservative 97; Mismatches 218; Indels 76; Gaps 16;

QY 11 SQAPSASALASESARPLADG-----LKSPLMKQAVRHHKKNLHNR 58
DB 113 SGAPPAAS-CGSSAAYSSSGRSHPTSGSSSHARSTGCGMSRSARANDDVHNGYK 171
QY 59 PLETLGKTYGKVKARE-SSGRVLAIKSRKDKIKDEODLLHIREIEMSLNHPHII 117
DB 172 LKTIQGNFPAKVLAAHILILGKEVAVKIIDKTAL-NPSLOKLFREVRIMKVLNHPHIV 230
QY 118 ALHEVFENSSKIVIMEYASRGDLYVIERPRLSERDARHPFQIYSALHYCHONGIYH 177
DB 231 KLYQVMEBQTLVLYLVEYASGGEVFDYLVAGRMKEKRAKRFQIVSAVQYLSKNITL 290
QY 178 RDLKLENIILDANGNIKADFGLSNLHKGFLOTFCSSPLVYASPEIYNGRPYGPEDSMS 237
DB 291 RDLKLENIILDODNMKIKADFGFNSLFFGKLDLTFCCSPYAAPBELFQGGKIDGPEVDV 350
QY 238 WSLGVLVLYLVHGTMPPDGDHKTLYKQISNGAYREPPKPSDAC-GLIRMLMVNPTBAT 296
DB 351 WSLGVLVLYLVSGSLPFDGQNLKELREYVLRGKTRIPFYMTDGCNLKFLIILNPSRG 410

[illegible]

```

RESULT 6
GB9287
Protein H39E23.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: GB9287
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.mucsl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: GB9287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 <STO>
C:Cross-references: GB:chr_V; PID:CAE09532.1; PID:G3878100; GSPDB:GN00023; CESP:H39E23
A:Gene: H39E23.1
A:Map position: 5

```

Query Match	20.3%;	Score 667;	DB 2;	Length 1246;
Best Local Similarity	31.9%;	Pred. No. 1e-19;		
Matches 182;	Conservative 93;	Mismatches 199;	Indels 96;	Gaps 16;

QY	40	KKQAVK-----RHHKHLRH-----RREFLETKGK 66
DB	10	KRRIVKVSPOGDHYTHNRKNSSSHARSTGSGMSSRSAARRNDODVHGKYLKLTITGKG 69
QY	67	TYGKVKKARE-SSGRLVAIKSIRKDKIKDEODLHIREIREIMSLNPHIILAEVEYN 125
DB	70	NPAVKVLAKAVITTHEVAIKIDITLAL-NBSSLQKLREVEIKMQLDHPNIVKLYQWET 128
QY	126	SSKIVIVEMEYASRGDLVDYIISERPLSERDARHFFROIVSALHYCHONGIYHRDLKENT 185
DB	129	EQLTYLVLEVASGGEVFDYLVAHGMRKEKARAFROIVSAVOYLHSNNIIHRLQKLENTL 188
QY	186	LIDANGIKIADLEGLSNLYHKGKFLQTFCCSPLYLASPEIYNGKPYVGDEVDSGLVILY 245
DB	189	LLDDDMNIKIADLFGRSNTFSLGNKLDTFCCGPPAAAPELFSGKTYDGEVDVMSLGIYLY 248
QY	246	ILVHGITMPFQODDKTLVKSISNGAYREPRPSDAC-GLIRWLMNVPTRATLEVDASH 304
DB	249	TLVSGSLPFPDQNLKEIRERVLKRTKRIPIFYMSDTCENLTKKFLVINIPQRRSSLDNIMKD 308
QY	305	WVWVMGYTTGVGEDEALREGGHPGSD-----FGASMDMLRRSSRPL 347
DB	309	RMNNVGY-----EDDELKPIEPPKQDIDEBIRIKLQIPQLGNNKALLESVKE--- 359
QY	348	LENGAKYCSFFPKQHAVPGGSGTVPGLEIROHSLIKRSKENDAAONLQDPAEDTSSRPGKSS 407

[illegible]

RESULT 7
T13741
hypothetical protein 22B5.8 - fruit fly (*Drosophila melanogaster*)
C|Species: *Drosophila melanogaster*
C|Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C|Accession: T13741
R|Murphy, L.; Harris, D.; Barrel, B.
submitted to the EMBL Data Library, April 1999
A|Description: Sequencing the discal X chromosome of *Drosophila melanogaster*.
A|Reference number: Z17668
A|Accession: T13741
A|Status: Preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-1398 <MUR>
A|Cross-references: UNIPROT:O77268; EMBL:AL031765; NID:e1371523; PID:e1329905; PTDN:CAA2
C|Genetics:
A|Cross-references: FlyBase:FBgn0000667
A|introns: 205/3; 227/1; 322/3; 688/3; 782/3; 814/2; 1363/3
A|Note: EG:22B5.8

	Query Match	20.2%;	Score 664;	DB 2;	Length 1398;	
	Best Local Similarity	31.1%;	Pred. No. 1.5e-19;			
	Matches	210;	Conservative	92;	Mismatches	216;
					Indels	158; Gaps
Oy	10	PSOAPSALASASARPLADGLIKSPKPYLMKKQAVKRHHNKNLHRRYFLETLGKTGTG	69			
Dd	110	PGPSTTSAVG-----AGGI--SGKULLIKAEWRMGF-----YDIERTIGKGNRA	153			
Oy	70	KVKAKRES-SGRVAIKSIRKDOKIDEODLLHIRREIEMSSLNHPIIAIHVEVENSRSK	128			
Dd	154	VKLAKRHRITKQNEVAIKTIIDSKQL-DQNLNGVYAREVEIMELTKGPRIHIKLQVMETKM	212			
Oy	129	IIVIMEVASRDVDVYSERPRLSRDRAHFRQCVSAHYCHONGIYVRDLKENILLD	188			
Dd	213	IYIVEVASOGEBIPYIAKYRMSSSARAFKFWQIIISAVECHHKGIYVRDLKAENILLD	272			
Oy	189	ANGNIKIADFGLSLNDYHKGFLQTFCCGSPLYASPEIVANGRPVCGEVDWSMGLVILYL	248			
Dd	273	LNNMIKIADPFESNMFKFGELLATMCSSPPAAAPVEFGSKQTGTGEIDMLSGVLVLYLV	332			
Oy	249	HGTMPEFDGDHKTLYKQISNGAYRBPXPDAQC-GLIFWLMVNPTFRATLEDVASHWM-	306			
Dd	333	CGALPFDGSTIQSLRDRLSGRFRIPIEFMSSECEHLIRRMVLFBPTRRYTDIOIKRHWM	392			
Oy	307	-----VMMGYTTGVGEQALNRGCHSPDFFGASAMDWI-----	340			
Dd	393	CPELLHEVLAKTNLGAERQRTSV---EPSEDILRI-MAEYVGJGSDKTRASLKKNYYDH	447			
Oy	341	-----RRSRPILLENG-----AKVCSFFKOHVPG-----GG	366			
Dd	448	VAAIYLLLDQDVSHHKKEQSNGLGASALTASTASRMIVSSRNDDQPTQQSQQQSKTIST	507			
Oy	367	STVCGLEGRHS-----LKSXKENDMAQNLOQDPABEDTSSRPGCKSLKP-KGILKK	417			
Dd	508	SSIIAKQOCHRLSRHQTVLMSERNNAHGAFTVDDPBGYYAKGPIQLPLPLTGHSHTL	567			
Oy	418	KSTSSGVEDQ-----POLLRVPDTPGQPVPA-VSLPRKITLKKSRQREGGYSS	469			

Db 568 TGYLNGGVGVNDASGILPLMRTPLP-TPASPAFNSCSTSRVG-----RHSLSSSS 619

Qy 470 P-----EPSESG-----ELLDASDVFGSDPEQKSPQASGLLHR---KGI 508

Db 620 PSHHRPAVLSLSDNNPSLANRCRMMA-----CGGVGAVGVPLAKQHLQITSEI 674

Qy 509 LKLNKFSRTALEGT-----DSTGSLDQLASSHPARSPS-----GA 549

Db 675 IQQSTEDCCALLQGGTAVAAGKDDPPKASSVGVGPVPASTPTSTSTAGPESGAPCPGE 734

Qy 550 VSEDSI---LSESPD 562

Db 735 INKXIKTKMSSSSSSFD 750

RESULT 8

T52633
serine/threonine-specific protein kinase (EC 2.7.1.1-) AKIN11 [validated] - Arabidopsis t
N/Alternate names: SNF1 protein kinase omo1og AKIN11
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #ext_change 09-Jul-2004
C/Accession: T52633
R/Bhalero, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machid
Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999
A/Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein ki
A/Reference number: Z25116; MUID:99238528; PMID:10220464
A/Accession: T52633
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-512 <BHA>
A/Cross-references: UNIPROT:P92958; EMBL:X99279; PIDN:CAA67671.1
A/Experimental source: cultivar Columbia
A/Genes: AKIN11
C/Function:
A/Description: EC 2.7.1.1-; serine/threonine-specific protein kinase AKIN11 [validated, M
Complements SNF1 mutations in yeast
C/Superfamily: AMP-activated protein kinase; protein kinase homolog
C/Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase

Query Match 19.7%; Score 649.5; DB 2; Length 512;
Best Local Similarity 38.6%; Pred. No. 2.2e-19;
Matches 152; Conservative 65; Mismatches 126; Indels 51; Gaps 9;

Qy 57 YEFLETLGKGYKVKARE-SSGRVAKSIKDKIKDEQDLHRRREIIMSSLNPH 115

Db 20 YLGGKTLGGSGFKVLAHVYVGHKVAIKILNRKIKMMEBEKVRRIKILRLFMHP 79

Qy 116 IIAHVEPENSRIIVIMEYASRGDIYDISSRPILSERDARHFFROIYSALHYCHONGI 175

Db 80 IIRQYEVIEITTDIYVMEYVSGELFDYIVKGRLODEBANKFPQOIIISGVEYCHRMNV 139

Qy 176 VHRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIIVNGKPYGPEV 235

Db 140 VHRDLKLENIILDSRCNIKIADFGLSNVARDGHFLTKTSCGSPVYAPAVISGKLYAGPBY 199

Qy 226 DWSMISGVLLYLIVHGMPPDGDHKTLYVQISNGAVREPPK-PSDCCGILRLMLMNPFR 294

Db 200 DWSMISGVLLYVALLCGLPDPDENIPNLFKKIKGGITLPSHSSEKRDILPRLIYDPVK 259

Qy 295 RATLESDVASHMWNMGYTTGVEGEALRGCGHPSGDFGRASMDMLRSGRPILLENGAVY 354

Db 260 RTIIEIRIHRWF-----QTHLPRIYAVPPDIVEAKKI 294

Qy 355 CSFFQHVPGGSGTYVGLERQSLK--KSRKENDMAQNTQGPDAEDTSSRPKSSILKPK 412

Db 295 NEEIYQEVVN-----MGFDRNQVLESLRRRTQND-----ATVTYLLILNDRRVPS 340

Qy 413 GILKK--KSSTSGVEQDEPQELRPVDPPIPGQPV 444

Db 341 GYLSSEFQETTDG---SNPMR---TPKAGASPV 368

RESULT 9

A26030
serine/threonine-specific protein kinase (EC 2.7.1.1-) SNF1 - yeast (Saccharomyces cerev
N/Alternate names: protein YBR477W
C/Species: Saccharomyces cerevisiae
C/Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #ext_change 09-Jul-2004
C/Accession: A26030; S69644
R/Celenza, J.L.; Carlson, M.
Science 233, 1175-1180, 1986
A/Title: A yeast gene that is essential for release from glucose repression encodes a p
A/Reference number: A26030; MUID:86289463; PMID:3526554
A/Accession: A26030
A/Molecule type: DNA
A/Residues: 1-633 <CEL>
A/Cross-references: UNIPROT:P06782; EMBL:M13971; NID:g172629; PIDN:AAA5058.1; PID:g1726
R/Dierich, F.S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A/Reference number: S69554
A/Accession: S69644
A/Molecule type: DNA
A/Residues: 1-633 <DIE>
A/Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64904.1; PID:g927732; GSPDB:GN000
C/Genetics:
A/Genes: SGD:SNF1; MIPS:YBR477W
A/Cross-references: SGD:S0002885; MIPS:YBR477W
A/Map position: 4R
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A/Note: required for expression of glucose-repressed genes in response to glucose depriv
C/Superfamily: AMP-activated protein kinase; protein kinase homolog
C/Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransf
F:61-69/Region: protein kinase ATP-binding motif
F:84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted
F:182,186/Binding site: magnesium (Asn, Asp) #status predicted
F:210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predict.

Query Match 19.7%; Score 648.5; DB 1; Length 633;
Best Local Similarity 33.1%; Pred. No. 2.9e-19;
Matches 161; Conservative 79; Mismatches 133; Indels 113; Gaps 14;

Qy 47 HHHKINLRH-----RYEFLTLGKGYKVKKA-RESSGRVY 82

Db 22 HHHHHHHHGGGSSSTLNNPKSLADGHNIGYQVKTLLGGSSFKVLAHYTTTGGQV 81

Qy 83 AIKSRKDKIKDEQDL-HIRREIIMSSLNPHITAHVEPENSRIIVIMEYASRGDI 141

Db 82 ALKIINK-KVLAKSDMQGRIEREIEIYLRILRHPHIILKYDVIKSKDEIIMVIEYAG-NEL 139

Qy 142 VYISERPRLSRDDARHPRROIYSALHYCHONGIYHRDLKLENIILDANGNIKIADFGLS 201

Db 140 FDIYIVQDRDMSQEARFFQOIIISAVEYCHRRKIYHRDLKLENIILDEHLNVAIKADFGLS 199

Qy 202 NLYHKGFLOTFCGSPLYASPEIIVNGKPYGPEVDSWISGLVLYLVHGMPPDGDHKT 261

Db 200 NIMTDGNFLTKTSCGSPVYAPAVISGKLYAGVEVDWMSGVLLYVWLCCRLPDDSEIYV 259

Qy 262 LVQISNGAYREPPKRS-DACGILRLMLMNPTRATLEVDVASHMWNMGYTTGVEGEQA 320

Db 260 LFKNISNGVYTLPKFLSPAGLIKRLMLIVNPLNRIISHIEMQD----- 303

Qy 321 LREGHPSGDFGRASMDMLRSGRPILLENGAKVCSFFQHVPGGSGTYVGLERQSLK 380

Db 304 -----DW-----FVVDLP-----EVLLEP 317

Qy 381 SRKENDMAQNTQGPDAEDTSSRPKSSILKPKGILKKSSSTSGVEQDEPQELRPVDPPI 440

Db 318 DLKPRPEENENNDSKKQSS---PDNDEIDNNIVMILSTNGYKDELYESLESSEDP 374

Qy 441 --GQVPVAVSL-PRKGIIKKSRQRESGYSSPPESEGEILLDASGVFVSGDP---VQ 493

Db 375 AFNEIRDAVMIKENKSLIKDMANKS-----VSDLEDTFLISQSPPTPQQS 421

Qy 494 KSPQAS 499
Db 422 KSHQKS 427

RESULT 10

gik protein - chicken
JC7500
N.Alternate names: Qtn-induced kinase
C.Species: Gallus gallus (chicken)
C.Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C.Accession: J07500
R.Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.
Biochem. Biophys. Res. Commun. 276: 564-570, 2000
A.Title: The new serine-threonine kinase, Qik, is a target of the qtn oncogene.
A.Reference number: J07500
A.Contents: Embryo fibroblasts
A.Accession: J07500
A.Molecule type: mRNA
A.Residues: 1-798 <Xia>
A.Cross-references: UNIPROT:Q91A88; GB:AF219232
C.Description: This protein, a member of the AMPK/SNFI family of serine/threonine kinases,
C.Keywords: protein kinase
A.Gene: qik
C.Keywords: protein kinase

Query Match 19.7%; Score 648; DB 2; Length 798;
Best Local Similarity 29.4%; Pred. No. 3.8e-19;
Matches 195; Conservative 94; Mismatches 223; Indels 152; Gaps 17;

11 SOAPASALASASAPLADGLIKSPKPLMKQAVKHHKHLRHRYEFLTLTGKTYGK 70
5 SEDASVPAPSAQPPRLVGF-----YDIERTLGKGNFV 39
71 VKKARESSGRV-VAIKSIRKDKIKDEODLHTRREIEIMSNPHIILHEVFENSXKI 129
40 VKLAHRYATKQVAKIKIDKTRL-DPSNLEKTYREVIQKLNHNHIIILKYVMERKML 98
130 VIWMEYASRGDIYDIISERPLSERDARHFPRQIYSALHYCHONGIYVBDLKLNLDA 189
99 YIVTFEPAKNGEMFDHLSNGHLSSEBARCKFWQIISAVEYCHSHHIVHDLKTENLLDA 158
190 NGNITADFGSLNHLHKKELOTFCGSPRYASPELVNGKPYVGVDSLSGLVLIYVH 249
159 NNNTKADFGKGNFYKSGEPLSTWCGSPYAAPEVGEYEGPHLDIWSLGVLYVLC 218
250 GTMPFGODHKTLYVQISNGAYREPKPSDAC-GILRWLMTNPTFRATLEDAVASHMWN 308
219 GSLPFDGNNLPTLRQVLEGRFRIRYFMSDECEITLIRLVVDPTKRITISQIKQHKM- 277
309 WGYTTGVEQELAREGHPGSDFGASMAWMLRRSRPL--LENGAKVCSFFKQHVPGG 365
278 -----QADPSLRQOQSLSFQMNVNSNLGADYNEQVLGI 310
366 GSTVGLERHSLIKK-----SRKEMDAQVLQODPAEDTSSRPCKSL 408
311 MOTTL-GIDRQRTVESLQNSSTNNHPALITYLLERLKEYSQLSRRPATGRQQRSSSTI 369
409 ---KLPKGIKKKSSSTSGEVEDPQELRPV-----PDPGQPV-----PAVSLPR 452
370 SNAEMRQDLSLSETLRSSILYQPOSILQPSIQAMDCMNNPDLQVFPVPVDPNENGFLR 429
453 KGLIKSFQREBSGYSSPEPSESGLLDASDVFSGDVPEQKSPQASGILLRKGILKLN 512
430 NRSISPSLSLETTL--SEVRQEKLED--EIKAVDHPIRIPS-----N 469
Qy 513 GKFEFTALEGTTPTSFGLSDQLASHPARPRPGAVSDESLTSSGFTDQ----- 563
Db 470 TSRHTTAAEVTHFYQHAAPCTVISSAS---PTEGTSDBCLTSSSSNDSSVALSSCIA 525
Qy 564 -----LDLPERLPETPLRGCVSVDNL-----GLLEQPPSEGIKR 597

Db 526 GQWMTGPATAMTSAFLASQSDAPVLQVQGMGASLLPVSFQEGRRASDTSLTQGLKA 585
Qy 598 WMQSE 601
Db 586 FROQ 589

RESULT 11

probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
N.Alternate names: protein kinase SNFI-related protein kinase
C.Species: Cucumis sativus (cucumber)
C.Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C.Accession: T10449
R.Stumpel, N.J.
submitted to the EMBL Data Library, December 1996
A.Reference number: Z17020
A.Accession: T10449
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-504 <GUN>
A.Cross-references: UNIPROT:P93113; EMBL:Y10036
A.Experimental source: cv. Masterpiece; cotyledon
C.Function:
A.Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C.Keywords: AMP-activated protein kinase; protein kinase homology
C.Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase
F.6-260/Domain: protein kinase homology <KIN>

Query Match 19.5%; Score 641; DB 2; Length 504;
Best Local Similarity 50.4%; Pred. No. 4.7e-19;
Matches 127; Conservative 47; Mismatches 76; Indels 2; Gaps 2;

57 YEFLETLGKTYGKYKARSS-SGRVLAIKSIRKDKIKDEODLHTRREIEIMSNPH 115
8 YKLGKTLGIGSGKVKIAHVALTGKVAIKINRRKIKNLDWEKVRREIKILRLFMHPH 67
116 IIAIHVEVENSXKIYIWEYASRGDIYDIISERPLSERDARHFPRQIYSALHYCHONGI 175
68 IIRLYEVEIETPSDIYVMEYVSGELFPYIYKGLQEBEAPNFQOIIISGVYCHRMNV 127
176 VHRDLKLNILLDANGNKIKIDPGLSNLYHKKPLQTCGSPLYVSPRIVNGKPYVGPV 235
128 VHRDLKPNLLDSKCNKIKIDPGLSNLMRDHFLTKCGSPNVAPEVTSKLYAGPEV 187
Qy 236 DWSLGLVLLIYLVHGTMPFGODHKTLYVQISNGAYREPKP-PSDACGLIRWLMVNPTR 294
Db 188 DWSGCVILLYALLCGTLFFDDENIPNLPRKIKGIGITLPSHLSGARELIPSLVYDPMK 247
Qy 295 RATEDVASHMW 306
Db 248 RITTEIRQHPW 259

RESULT 12

serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N.Alternate names: protein kinase SNFI homolog
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C.Accession: U01446; S58266; S66334
R.Luegten, L.; Thomas, M.; Blanchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A.Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A.Reference number: U01446; MUD:99013041; PMID:1339373
A.Accession: U01446
A.Molecule type: DNA
A.Residues: 1-512 <LEG>
A.Cross-references: UNIPROT:Q38997; GB:M93023; NID:9166599; PIDN:AAA2736.1; PID:9166600
R.Thiemmer, F.; Kirchner, M.; Teuber, R.; Dietrich, P.
submitted to the EMBL Data Library, May 1995
A.Description: Differential accumulation of the transcripts of 22 novel protein kinase g

A:Reference number: A56009; MWID:94217693; PMID:8164654
 A:Accession: A56009
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-511 <MOR>
 A:Cross-references: UNIPROT:Q40544; GB:D26602; NID:g946384; PIDN:BA05649.1; PID:g946384
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate
 C:Superfamily: AMP-activated protein kinase; protein kinase homologous
 C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
 F:17-271/Domain: protein kinase homology <KIN>
 F:25-33/Region: protein kinase ATP-binding motif
 F:148,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
 F:147,151/Binding site: magnesium (Asn, Asp) #status predicted
 Query Match 19.2%; Score 631; DB 1; Length 511;
 Best Local Similarity 49.6%; Pred. No. 1,2e-18;
 Matches 125; Conservative 48; Mismatches 77; Indels 2; Gaps 2;
 QY 57 YELELETLGKGTGYKKARRS-SGRVLAISIRKDKIKEDQDLHRRREIEMSSLNHPH 115
 DB 19 YKLGKTLIGSPFKVKAIEHTLLTGHKVAVKIKRRRIKNNEMEKYRRRIKILRLFNHPH 78
 QY 116 IIAIHVEFENSXIVIVMEYASRGDLVDYSESRPRKSESDARAFPFQIVALTAYCHONGI 175
 DB 79 IILYEVETPPSDIIVMEYKSGELFDYIVYEGRIQEDBARKEFQOIIISGVYCHRMV 138
 QY 176 VHRDLKLENTILDANGNIKADFGLSNLYHKGFLOTFCGSPLYASPEIYNGRPYGVPEV 235
 DB 139 VHRDLKPEMLLDSSKNVVKADFGLSNIMRMDGHFLMTSCGSPYVAPEVISGLYAGPEV 198
 QY 236 DSNLSGLVLLIYVHGMPDPGDQHKTLVKQISNGAYREPKE9-DACGLIRMLMTNPTR 294
 DB 199 DVMSCGIIVALLCGTLPPDENIPLYFKKIKGSMISLPESHLSAGARDILPRMLIYDPMK 258
 QY 295 RATLEDVASHHW 306
 DB 259 RMTIPETRMHP 270
 RESULT 14
 SS1025
 [hydroxymethylglutaryl-CoA reductase (NADPH2)] kinase (EC 2.7.1.109) - human
 N:Alternate names: AMPK
 N:Contents: [acetyl-CoA carboxylase] kinase (EC 2.7.1.128); [hydroxymethylglutaryl-CoA
 C:Species: Homo sapiens (man)
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C:Accession: SS1025; I38503
 R:Berl, R.K.; Matley, A.E.; See, C.G.; Sowith, W.F.; Aguan, K.; Garling, D.; Scott, J.
 FEBS Lett. 356, 117-121, 1994
 A:Title: Molecular cloning, expression and chromosomal localization of human AMP-activated
 A:Reference number: SS1025; MWID:95080410; PMID:7988703
 A:Accession: SS1025
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-552 <BER>
 A:Cross-references: UNIPROT:P54646
 R:Aguan, K.; Scott, J.; See, C.G.; Sarkar, N.H.
 Gene 149, 345-350, 1994
 A:Title: Characterization and chromosomal localization of the human homologue of a rat
 A:Reference number: I38503; MWID:95047501; PMID:7959015
 A:Accession: I38503
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-179, 'T', 181-270, 'G', 272-402, 'RO', 405-552 <RES>
 A:Cross-references: EMBL:U06454; NID:g758366; PIDN:AAA64745.1; PID:g758367
 C:Genetics:
 A:Gene: GDB:PRKAA2, PRKAA, AMPK, hAMPK
 A:Cross-references: GDB:451905; OMIM:600497
 A:Map position: 1p31-1p31
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-
 A:Note: in vivo substrates include hydroxymethylglutaryl-CoA reductase (NADPH) and acetyl

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:13:49 ; Search time 44.2557 Seconds
(without alignments)
1064.351 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293
Sequence: 1 MESVALLQRPQAPSASALA.....DQEVYAAVYALGICSKLS 631Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	100.0	631	US-09-579-664B-11	Sequence 11, App1
2	3293	100.0	631	US-10-355-975A-11	Sequence 11, App1
3	3278.5	99.6	630	US-10-355-975A-38	Sequence 38, App1
4	703.5	21.4	729	US-08-677-298-2	Sequence 2, App1
5	703.5	21.4	729	US-09-523-849-33	Sequence 33, App1
6	700.5	21.3	602	US-09-949-016-7417	Sequence 7417, App1
7	700.5	21.3	602	US-09-949-016-7418	Sequence 7418, App1
8	698.5	21.2	724	US-09-949-016-8234	Sequence 8234, App1
9	694.5	21.1	713	US-09-538-092-1022	Sequence 1022, App1
10	694.5	21.1	713	US-09-949-016-6214	Sequence 6214, App1
11	691.5	21.0	793	US-09-523-849-32	Sequence 32, App1
12	688	20.9	745	US-09-523-849-36	Sequence 36, App1
13	680	20.6	724	US-09-984-890-2	Sequence 2, App1
14	680	20.6	724	US-10-274-194-2	Sequence 2, App1
15	678.5	20.6	691	US-09-949-016-8255	Sequence 8255, App1
16	678.5	20.6	691	US-09-949-016-8256	Sequence 8256, App1
17	677.5	20.6	722	US-09-984-890-4	Sequence 4, App1
18	677.5	20.6	722	US-10-274-194-4	Sequence 4, App1
19	675	20.5	779	US-08-817-832B-31	Sequence 31, App1
20	673.5	20.5	722	US-08-817-832B-32	Sequence 32, App1
21	653	19.8	604	US-09-523-849-35	Sequence 35, App1
22	648.5	19.7	633	US-08-557-006C-43	Sequence 43, App1
23	648.5	19.7	633	US-09-538-092-212	Sequence 212, App1
24	648.5	19.7	633	US-09-633-328B-3	Sequence 3, App1
25	648.5	19.7	633	US-09-824-735-3	Sequence 3, App1
26	648.5	19.7	1203	US-09-799-875-5	Sequence 5, App1
27	645.5	19.6	512	US-09-633-328B-2	Sequence 2, App1

28	642	19.5	557	4	US-09-949-016-10174	Sequence 10174, App1
29	636.5	19.3	552	4	US-09-824-735-4	Sequence 4, App1
30	635.5	19.3	345	3	US-09-101-146-1	Sequence 1, App1
31	635.5	19.3	716	4	US-09-523-849-34	Sequence 34, App1
32	631	19.2	511	4	US-09-633-328B-4	Sequence 4, App1
33	628.5	19.1	552	4	US-09-538-092-1212	Sequence 1212, App1
34	625.5	19.0	418	4	US-09-248-796A-18441	Sequence 18441, App1
35	625	19.0	257	3	US-09-101-146-6	Sequence 6, App1
36	623.5	18.9	552	3	US-08-557-006C-40	Sequence 40, App1
37	619.5	18.8	257	2	US-07-857-224B-25	Sequence 25, App1
38	613.5	18.6	778	4	US-10-116-326-2	Sequence 2, App1
39	613.5	18.6	778	4	US-10-003-690-2	Sequence 2, App1
40	599.5	18.2	668	4	US-09-930-181-2	Sequence 2, App1
41	589.5	17.9	521	4	US-09-949-016-7570	Sequence 7570, App1
42	549	16.7	1064	4	US-09-538-092-154	Sequence 154, App1
43	547	16.6	703	4	US-10-116-326-6	Sequence 6, App1
44	528.5	15.0	603	4	US-09-930-181-17	Sequence 17, App1
45	525	15.9	260	2	US-07-857-224B-28	Sequence 28, App1

ALIGNMENTS

RESULT 1

US-09-579-664B-11
; Sequence 11, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OR INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-579-664B-11

Query Match	Score 3293; DB 4;	Length 631;
Best Local Similarity 100.0%;	Pred. No. 2.4e-277;	Indels 0; Gaps 0;
Matches 631; Conservative	0; Mismatches	0; Indels
QY	1	MESVALLQRPQAPSASALASARPLADGLTSPKPLMKQAVKHHKHLRHRVEFL 60
DB	1	MESVALLQRPQAPSASALASARPLADGLTSPKPLMKQAVKHHKHLRHRVEFL 60
QY	61	ETLGGKTYGVKKARSSGRVAIKSIRKDKDEBDLHIREIRIEMSLNPHIAT 120
DB	61	ETLGGKTYGVKKARSSGRVAIKSIRKDKDEBDLHIREIRIEMSLNPHIAT 120
QY	121	EYFENSSKVIYWEYASRGDLIDYISERPLSRDRAHFRFQVSLAHYCHONGIYHRDL 180
DB	121	EYFENSSKVIYWEYASRGDLIDYISERPLSRDRAHFRFQVSLAHYCHONGIYHRDL 180
QY	181	KLENTILDANGNINKINDPGLSNLYHKGFLQPCGSLVASPIYNGKPYGEVDSWSL 240
DB	181	KLENTILDANGNINKINDPGLSNLYHKGFLQPCGSLVASPIYNGKPYGEVDSWSL 240
QY	241	GVLLYLIVLWGTMPDQDHTLVKQISNGAYRPPPSDAGCIRWLLMVPTRATLBD 300
DB	241	GVLLYLIVLWGTMPDQDHTLVKQISNGAYRPPPSDAGCIRWLLMVPTRATLBD 300
QY	301	VASHMVMVNGYTTGVEQERLREGRHPSGDFGASADWLRRSSPLLENAGAVCSFFQ 360
DB	301	VASHMVMVNGYTTGVEQERLREGRHPSGDFGASADWLRRSSPLLENAGAVCSFFQ 360

QY 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 420
| | | | |
DB 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 420
QY 421 TSSGEVQEDPOELRPVDPDPGQVPAVSLPRKGIILKSKRQRESGYSSPEPESEGLD 480
| | | | |
DB 421 TSSGEVQEDPOELRPVDPDPGQVPAVSLPRKGIILKSKRQRESGYSSPEPESEGLD 480
QY 481 ASDVFVSGDPVQKSPQASGLLHRRKGIILKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
| | | | |
DB 481 ASDVFVSGDPVQKSPQASGLLHRRKGIILKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
QY 541 ARPSRPSGAVSEDSIISSESPDOLDPERLPETPLRGCVSDNLRGLQPPSEGLKRMWQ 600
| | | | |
DB 541 ARPSRPSGAVSEDSIISSESPDOLDPERLPETPLRGCVSDNLRGLQPPSEGLKRMWQ 600
QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 2
US-10-355-975A-11
; Sequence 11, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355, 975A
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975A-11

Query Match 100.0%; Score 3293; DB 4; Length 631;
Best Local Similarity 100.0%; Pred. No. 2,4e-277;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESVALLQRPQAPASALASASARPLADGLIKSPKPLMKQAVKRRHHKHLRHRYYFL 60
| | | | |
DB 1 MESVALLQRPQAPASALASASARPLADGLIKSPKPLMKQAVKRRHHKHLRHRYYFL 60
QY 61 ETLGKGTGKVKKARESSGRVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAH 120
| | | | |
DB 61 ETLGKGTGKVKKARESSGRVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAH 120
QY 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
| | | | |
DB 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
QY 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
| | | | |
DB 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
QY 181 KLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVDSWSL 240
| | | | |
DB 181 KLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVDSWSL 240
QY 241 GVLVYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 300
| | | | |
DB 241 GVLVYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 300
QY 301 VASHHVMVWNGYTTTGVEQELAREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ 360
| | | | |
DB 301 VASHHVMVWNGYTTTGVEQELAREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ 360
QY 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 420
| | | | |

DB 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 420
| | | | |
QY 421 TSSGEVQEDPOELRPVDPDPGQVPAVSLPRKGIILKSKRQRESGYSSPEPESEGLD 480
| | | | |
DB 421 TSSGEVQEDPOELRPVDPDPGQVPAVSLPRKGIILKSKRQRESGYSSPEPESEGLD 480
QY 481 ASDVFVSGDPVQKSPQASGLLHRRKGIILKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
| | | | |
DB 481 ASDVFVSGDPVQKSPQASGLLHRRKGIILKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
QY 541 ARPSRPSGAVSEDSIISSESPDOLDPERLPETPLRGCVSDNLRGLQPPSEGLKRMWQ 600
| | | | |
DB 541 ARPSRPSGAVSEDSIISSESPDOLDPERLPETPLRGCVSDNLRGLQPPSEGLKRMWQ 600
QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 3
US-10-355-975A-38
; Sequence 38, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355, 975A
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 38
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975A-38

Query Match 99.6%; Score 3278.5; DB 4; Length 630;
Best Local Similarity 99.8%; Pred. No. 4,4e-276;
Matches 630; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MESVALLQRPQAPASALASASARPLADGLIKSPKPLMKQAVKRRHHKHLRHRYYFL 60
| | | | |
DB 1 MESVALLQRPQAPASALASASARPLADGLIKSPKPLMKQAVKRRHHKHLRHRYYFL 60
QY 61 ETLGKGTGKVKKARESSGRVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAH 120
| | | | |
DB 61 ETLGKGTGKVKKARESSGRVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAH 120
QY 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
| | | | |
DB 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
QY 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
| | | | |
DB 121 EYFENSSKVIIVMEYASRDLDYISERPRLSERDARHFRQIVSALHCHONGIVHRDL 180
QY 181 KLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVDSWSL 240
| | | | |
DB 181 KLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVDSWSL 240
QY 241 GVLVYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 300
| | | | |
DB 241 GVLVYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 300
QY 301 VASHHVMVWNGYTTTGVEQELAREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ 360
| | | | |
DB 301 VASHHVMVWNGYTTTGVEQELAREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ 360
QY 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 420
| | | | |
DB 361 HYPGGGTYVGLERHSLKSKRKENDMAQNLQGDPAEDTSSRPKGSILKLPKGIILKKSS 419

QY 421 TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGLKKSRQRRSGYSSPESSEGLD 480
DB 420 TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGLKKSRQRRSGYSSPESSEGLD 479
QY 481 ASVPFSGDVEQKSPQASGLLHRKGLTLKNGKFSRTALEGTTPTFGSLDOLASSHPA 540
DB 480 ASVPFSGDVEQKSPQASGLLHRKGLTLKNGKFSRTALEGTTPTFGSLDOLASSHPA 539
QY 541 ARPSRSGAVSEDSILSSSEFDDLDLPERLPETPLRCGVVDNLRLGELQPPSEGLKRMQ 600
DB 540 ARPSRSGAVSEDSILSSSEFDDLDLPERLPETPLRCGVVDNLRLGELQPPSEGLKRMQ 599
QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 600 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 630

RESULT 4

US-08-677-298-2
; Sequence 2, Application US/08677298
; Patent No. 5863729
; GENERAL INFORMATION:
; APPLICANT: Pivnicka-Worme, Helen
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAP-1
; TITLE OF INVENTION: KINASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,298
; FILING DATE: 09-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 9-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-677-298-2

Query Match 21.4%; Score 703.5; DB 2; Length 729;
Best Local Similarity 35.1%; Pred. No. 3.7e-52;
Matches 195; Conservative 76; Mismatches 180; Indels 105; Gaps 18;
QY 57 YEFLETLGKGTGYKVKARE-SSGRVVAIKSIRKDKIKDQDILLHIREIINSLNHPH 115
DB 56 YRLKTIIGKNFPAKVLARHILTGREVAIKIIDKTQL-NPTSLQKLFREVRIMKILNHPN 114
QY 116 IIAIHEVFENSKIIVIMEYASRGDLVDYISERPRLSERDARHPFROIYVSAHYCHONGI 175
DB 115 IYVLFVEIETEKTLIYIMEYASGGEVFDYLVAHGMRMEKEARSKFRQIVSAVOYCHQKRI 174
QY 176 VHRDLKLENTLLANGIKITADRGSLNLYHKGKLFOTFGCSPLYASFEIYNGKPYVPEV 235
DB 175 VHRDLKLENTLLADNMKIKIADFGFSNEFTVGGKLDTFGCSPPYAABELFOGKKYDGPV 234

QY 236 DSWSLGVLYIYLVHGTMPDGDHKTLYVQISNGAYREPKPSDAC-GLIRMLMNPTR 294
DB 235 DWLSLGVLYIYLVHGTMPDGDHKTLYVQISNGAYREPKPSDAC-GLIRMLMNPTR 294
QY 295 RATLEPVAASHMVMVWGYTTCVGEQALRECGHPSGDFGRASMDMLRRSSRPILLENGAKV 354
DB 295 RGTLEQIMKDRWNNAH-----BEDLKPVEBELDISQKRLDIM-----V 336
QY 355 CSFFKQHVPGGSGTVGELERQHSKKSRKENDMAQNL---QGDPAEDTSSRPKSSIKLP 411
DB 337 GNGVSGE-----ELQESLSMKKYDELTATYLLGLKSKSELDAOSSSSSNLSLA 385
QY 412 KGIKKKSSSTSGE-----VQ---EDPQLRPVDPTRGQVPAVSLPRK-----GIL 456
DB 386 KVPSSDLNNSTQSGPHHKVQBSVSSQKORRYSDHAGPAIPSVVAYPKRSQRTADGDL 445
QY 457 KK---SRQRES-----GYSSPESSEGLDASDFVSGDVEQKSPQ 497
DB 446 KEDGSSRSKSSGAVGKGIAAPASPMLGANASNPKADIPERKKSSTV-----PSSNT 497
QY 498 ASGLLHRK-----GILKNGKFSRTALEGTTPTFGSLDOLASSH---PAA 541
DB 498 ASGGMTRNTYVCSERTTADRHSHVIO-NGKENSTIIDQRTF-----VASTHSISSAA 548
QY 542 RPSR---PSGAVSEDS 554
DB 549 TPDRIIRPRTAIRST 564

RESULT 5

US-09-523-849-33
; Sequence 33, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank Accession No. 6458561 g3089349
US-09-523-849-33

Query Match 21.4%; Score 703.5; DB 4; Length 729;
Best Local Similarity 35.1%; Pred. No. 3.7e-52;
Matches 195; Conservative 76; Mismatches 180; Indels 105; Gaps 18;
QY 57 YEFLETLGKGTGYKVKARE-SSGRVVAIKSIRKDKIKDQDILLHIREIINSLNHPH 115
DB 56 YRLKTIIGKNFPAKVLARHILTGREVAIKIIDKTQL-NPTSLQKLFREVRIMKILNHPN 114
QY 116 IIAIHEVFENSKIIVIMEYASRGDLVDYISERPRLSERDARHPFROIYVSAHYCHONGI 175
DB 115 IYVLFVEIETEKTLIYIMEYASGGEVFDYLVAHGMRMEKEARSKFRQIVSAVOYCHQKRI 174
QY 176 VHRDLKLENTLLANGIKITADRGSLNLYHKGKLFOTFGCSPLYASFEIYNGKPYVPEV 235
DB 175 VHRDLKLENTLLADNMKIKIADFGFSNEFTVGGKLDTFGCSPPYAABELFOGKKYDGPV 234
QY 236 DSWSLGVLYIYLVHGTMPDGDHKTLYVQISNGAYREPKPSDAC-GLIRMLMNPTR 294

```
Db      235 DWASLGLVILYTLVSGSLPFGCGNLKELERVLGAKTRIPFYMTSDCENLKRFLVLANPIK 294
Qy      295 RALTEVASHWVWVWNGYTTGVGEQALRECGHPSGDFGRASMDWLRSSRPILLENGAKV 354
Db      295 RGLTEIMKDRMINAGH-----EDELKPFVEPELDISDQKRIDIM-----Y 336
Qy      355 CSFEEKHVPGGSGTVGLERQHLKSKRKENMDMAQNL---QGDPAETSSRPCKSLKLP 411
Db      337 GNGYSDB-----EIQESLSKMKYDEITRTYLLLRKSESLDASSSSSLSLA 385
Qy      412 KGLIKKSSSTSGE-----VQ---EDFQELRPVDPFGQVPVAVSLPRK-----GIL 456
Db      386 KVRPSSDLNNSTQSPHMKVQRASVSSQKQRRYSDDHGAIPSVVAPKRSQSTADGL 445
Qy      457 KR---SRQES-----GYSSPSPSESGELLDASDVFSVSGDPVEQKSPQ 497
Db      446 KEGISRSKSSGSAVGKGIAPASPMIGNASNENKADIPERKKSSTV-----PSSNT 497
Qy      498 ASGLLHRK-----GILKNGKFSRTALEGTTPTFGSLDQLASH---PAA 541
Db      498 ASGGMTTRRTYTCSERTYADRHSHVIO-NGKENSTIPDQRT-----VASTHSISSAA 548
Qy      542 RPSR---PSGAVSEDS 554
Db      549 TPDRIFRPRTASRST 564
```

```
RESULT 6
US-09-949-016-7417
; Sequence 7417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7417
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7417
```

```
Query Match      21.3%; Score 700.5; DB 4; Length 602;
Best Local Similarity 31.2%; Pred. No. 5e-52;
Matches 196; Conservative 100; Mismatches 216; Indels 117; Gaps 21;

Qy      23 SARPLADGLIKSPKPLMKQAVKRRHHKKNLRRH-----YEFLETLGKGYGVKKAR- 76
Db      34 SSRPTLGH--DSKSSKSNMTRGRNSATSADQPHIGNRLLKTTGKNFAVKLARHI 91
Qy      77 SSGRLVAISIRKDKIKDEODLHTRREIEMSLNHPHIIAHVEVFENSSKIIVMEYA 136
Db      92 LUGKEVAVKIIDTQI--NSSLOKLPREVRIMKVLNHPNIVKLPFVLETEKTLVMEYA 150
Qy      137 SRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIVHRDLKLENTILDANGIKIA 196
Db      151 SCGEVFDYLVAGHMKKEKRAKFRQIVSAVOYCHQKFTVHRDLKLENTILDANGIKIA 210
Qy      197 DFGLSNLVHKGKFLQTFGCSPLVASEIIVNGKPYVGPVSDWSLGLVLYLVHGTMPFDG 256
Db      211 DFGSNEFTFGNKLDPFGCSPPYAAPBELFGKKYDGPVSDWSLGLVLYLVSGSLPFDG 270
```

```
Qy      257 QDHKTLVKQISNAGVREPPKPSDAC--GLIRMLMNVPTRRATLEDVASHWVWNGYTTGV 315
Db      271 QNKELEREVLKRTKRIIPFYMTSDCENLKLPLILPSKRGTLQIMKDRMINVGH---- 336
Qy      316 GEQALRECGHPSGDFGRASMDWLRSSRPILLENGAKVCSFFQHVPGGSGTVPELQERO 375
Db      327 -EDELKPFVEPELDISDQKRIDIM-----Y 336
Qy      376 HSLIKSRKENMDA-----ONTQGD-----PAED-TSSRPCKSLKLPKGI----- 414
Db      362 DSLVGRYNEVWATYLLLYGKSSSELEGDTITLKPSPADLTNNSAPSPHKTQVRSVANP 421
Qy      415 -----LKKSSSTSGE---VQEDPQELRPVDP---PQGVPAVSL 450
Db      422 KORFSDQAPALPTSNYSYSKTQSNMANKRPEEDRESGRASSATKVPASPLP----- 476
Qy      451 PRGILKSGRORESGYSSPSPSESGELLDASDVFSVSGDPVOKSPQASGLLHRKGIK 510
Db      477 ---GLERK-----TTPTPS-TNSVLTSTNRSNPLERASLQASIQNGKDSLT 524
Qy      511 LNKFSRTALEGTTPTFGSLDQLASHPAARPSRPSGAVSBSISPSFQDLPERL 570
Db      525 MGSRASTASAGAASVAAARPHQKMSMSASVHPNKASG-----LPPTESNGEVPAPROV 578
Qy      571 PETPLRGYVVDNRLGLEQPPSEGLKRW 599
Db      579 PP---ACPCCLPPIRPHQ-----QQW 597
```

```
RESULT 7
US-09-949-016-7418
; Sequence 7418, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7418
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7418
```

```
Query Match      21.3%; Score 700.5; DB 4; Length 602;
Best Local Similarity 31.2%; Pred. No. 5e-52;
Matches 196; Conservative 100; Mismatches 216; Indels 117; Gaps 21;

Qy      23 SARPLADGLIKSPKPLMKQAVKRRHHKKNLRRH-----YEFLETLGKGYGVKKAR- 76
Db      34 SSRPTLGH--DSKSSKSNMTRGRNSATSADQPHIGNRLLKTTGKNFAVKLARHI 91
Qy      77 SSGRLVAISIRKDKIKDEODLHTRREIEMSLNHPHIIAHVEVFENSSKIIVMEYA 136
Db      92 LUGKEVAVKIIDTQI--NSSLOKLPREVRIMKVLNHPNIVKLPFVLETEKTLVMEYA 150
Qy      137 SRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIVHRDLKLENTILDANGIKIA 196
Db      151 SCGEVFDYLVAGHMKKEKRAKFRQIVSAVOYCHQKFTVHRDLKLENTILDANGIKIA 210
Qy      197 DFGLSNLVHKGKFLQTFGCSPLVASEIIVNGKPYVGPVSDWSLGLVLYLVHGTMPFDG 256
Db      211 DFGSNEFTFGNKLDPFGCSPPYAAPBELFGKKYDGPVSDWSLGLVLYLVSGSLPFDG 270
```



```

OY 245 RATTEDVASHMMVNMVYTTGVGOEQLRGGHSGGFGRA5MADWLRSSRPLENGAKV 354
Db 225 RGTLLQIMKDRIMNIGH-----EEDLKPFEVELDI5DQKRLDIM-----V 336
OY 335 CSFFQOHVPGGGSYVPLGEROHLKKSRENDMAONL-----QGDPAEDTSSRPKSS 407
Db 337 GNGVGOE-----EIOESL5TKKDEIATATALLGRKSSEVRPSSDLNNGOS- 384
OY 408 LKLPLGLKLLKKSSTSGEVOEDQELRPVDPTRGGQVPAVSLPR-----452
Db 385 ---PHHKVGRSVSS-----QKRRKRSDHAGPEIPSVVAFPK5QSTADSLKEDG 433
OY 453 -----KGLKKSRORESGYSSPEBSEGGELIDASDVFSGDVPVOKSPQAS 499
Db 434 ISSRKSSTGSAGVGKGIAPASPM--LONANPNVADIPERKKSSTV-----P5SNTAS 483
OY 500 GILLHARK-----GILTKNGKESRTALEGTPESTFSLDQLASH---BAAP 543
Db 484 GGMTRRNMYVCSERTTDDRH5VIG-NGKENSTIPDQRT-----VASTH5ISSAATP 534
OY 544 SR---PSGAVSEDS 554
Db 535 DRIRPRGTASRST 548

```

```

RESULT 10
US-09-949-016-6214
: Sequence 6214, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6214
: LENGTH: 713
: TYPE: PRT
: ORGANISM: Human
US-09-949-016-6214

```

Query Match	21.1%	Score 694.5	DB 4	length 713
Best Local Similarity	34.5%	Pred. No. 2.1e-51		
Matches	191	Conservative	75	Mismatches 171; Indels 117; Gaps 17;
QY	57	YEFLETLKGGTYGVKKARE-SGGRLVAIKSRKDKIDKEDDLHRRREIEMSSINHPH	115	
Db	56	YRLKTIKIGKSNPAKVLARHILVTGREVALIKIDKQL-NPTSLQKLFREVRIMKILNHEN	114	
QY	116	IIATHEVENSRSKIYIVMEYASRGDLVDYISERPRLSDDARHFFQIYVSALHYCHONGI	175	
Db	115	IVKFEVETOKTLVILIMEYASGGKFDVLVAHGRMKEKRAASKPFQIYSAVOYCHQKEI	174	
QY	176	VHRDLKLENILLDANGINKIADFGSLNLVHKGKFLDTFGSGSPLYASPELVNGCPVGPPEV	235	
Db	175	VHRDLKAENLLDADMNIKIADFGGSNFETTVGGKDLTFGSPRYAAPBELFGGKDYGPPEV	234	
QY	236	DSWSIGVLLYIIVHGTMPEFGQDHHKTLVVOISNGAYRBPXPSDAC-GLIRWLLMWNPTR	294	
Db	235	DWSIGVILTYLVSSGLPFDGQNLKELRBRVLRGKKRIRIFPMYSTDCENLLKRLVLYNPPIK	294	
QY	295	RATLEDVASHMWVNVNGYTTTGVGGEALREGGHPSGDFGASADWIRRSRPILENGARV	354	

```

Db      295  RGLLEGIIMDRWIMNGH-----EEDELKRFYVPELLEISDQKRIIM-----V 336
QY      355  CSFFPKQNVPEGGSIVYGLERQHSLLKSKKENDMAQL-----QGDAEPTSSRPKSS 407
Db      337  GWCYSQE-----EIQESLSKMTYDEITATAYLLGKSSGEVAPSSDLNNSTGOS- 384
QY      408  LKLPKGLKKKSSSTSGEVOEDPQELRVPPTPCQVPVAVSLPR-----452
Db      385  ---PHKVVQSRVSSS-----QKQRYSDBAGGIPSVVAVYFKRSQTSIAOSDLKDG 433
QY      453  -----KGLIKTSRQREGYIYSSPEPSSGSLDASIVPVSQDPVECKSKQAS 499
Db      434  ISSRKSTGSAVGCKGIAPASPM--LGNASNPKAKDIPERKKSIVY-----PSSVTAS 483
QY      500  GLDLHRR-----GILKLNKGFSTALGPTSPFGSLDOLAASH---PAAP 543
Db      484  GGMTRNTNYVCSERTTDDRHSHVIG--NGKENSTIPDQRT-----VAGSHSTSSAATP 534
QY      544  SR---PSGAVSEDS 554
Db      535  DRIRPFRGTASRST 548

```

```

RESULT 11
US-09-523-849-32
; Sequence 32, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bobotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURES:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank Accession No. 6458561_g20521899
US-09-523-849-32

```

	Query Match	21.0%; Score 691.5; DA 4;	Length 793;
	Best Local Similarity	35.6%; Pred. No. 4.6e-51;	
	Matches 187;	Conservative 76;	Mismatches 189; Indels 73; Gaps 14
Oy	57	YEFLETLGKGYGVKKAKARE--SSGRVAKISIRKDKIDEDDLHLHREIEIMSSLNHPH 115	
Dd	60	YLRLQTKTGKGNFAKVKLARHVLTGREVAVKIIDTQL--NPLSLQKLFRFVRIMKILNHPH 118	
Oy	116	IIAIHVFENSKSIYIVMEVASRGVLXYISERPRLSEPDARHPHFROIYVALHYCHONGI 175	
Dd	119	IYVLFVEIETKTLYLVMEIYASGSEVPFIYLAHGMRKEKERAKFROIYVANVOYCHOCKI 178	
Oy	176	VHRDLKLENILLDANGNIKIADFGLSNLVHKGLQTFCGSPLYASBEIVNGKPYVGPEV 235	
Dd	179	VHRDLVAEMILLDALMNIIKIADFPGSNFTFYGNLDFPCSGPFAAELFLQGKKYDGPEV 238	
Oy	236	DSMSGLVLLIYIVHGMTPFDGDHHTIKOISNAYABEPKPSPAC--GIIRLLLMNNPR 294	
Dd	239	DWMSGLVIITYTVSSSLPFDDGNLKELERLYRKRKYVPPPTMSTDCEMLKLLVLPPIK 298	
Oy	295	RATLEDVASHMMVNNGYITYGVEODALREGGHPSGDFGRASMDWLRRSRPLL--NGAK 353	
Dd	299	RGSLEGEIMDRMMNVNGH-----EEBELKPYSEPPLDLNDKRIIDINWTMFAPADEINDAL 353	


```

QY 437 PDT-----PCGPVAVNSLIPKPKILKTSRQRESGYSSSPESSESELLDASVFGSDVEQ 493
Db 447 SSTAKVPASPLP-----GLERKK-----TPTPTSNVLSTSTN-----RGR 483
QY 494 KSPQASGLLHKKILKLT---NGKSRRTALGCTPTSTSLDQALASHNPA--RPSRPSGA 549
Db 484 NSP-----LLEBASLGGQASLQNGKOS-TAPQ-RVPVASBSAHNITSSGGA PDTPTNPPRGV 536
QY 550 VSEBSILSS---ESFDQDLPERLL-PETP 574
Db 537 SSRSTFHAGQLRQVRDQQLLPYGVVPASP 565

```

```

RESULT 14
US-10-274-194-2
; Sequence 2, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306D1V
; CURRENT APPLICATION NUMBER: US/10/274,194
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-274-194-2

Query Match          20.6%; Score 680; DB 4; Length 724;
Best Local Similarity 32.0%; Pred. No. 4e-50;
Matches 201; Conservative 99; Mismatches 195; Indels 134; Gaps 25;

QY      6  L L Q R P S Q A P S A L A S E S A R P L A D G L K S P Y P L M K Q A V K H H K H N L R H E --- Y E F L 60
Db      11 L N E R D T R O P T L G H I N D S ----- K P S S K S M I R G R N A T S A D E P H I G N R L L 56

QY      61 E T L K G Y T G V K Y K A R E - S S G R L V A I K S I R K O K I D E Q D L H I R R E I E M S L A N P H I T A I 119
Db      57 K T I G K G N F A K K A L A R H I L T G E V A V K I I D K T Q L - N S S L Q L C F R E V R I M V K L A N P N I V K L 115

QY      120 H E V E N S S K I V I N E Y A S R G D L Y D I S E R P L S R D A N H F P R Q I V S A L H Y C H Q N G I Y H R D 179
Db      116 F E V I E T E R K T I L V M E Y A S G E V F P D Y L V A H G M K E K R A K F R Q V A S A V Q C H Q F I Y H R D 175

QY      180 L K E N I I L L D N G N I K I A D F G L S N L Y H K G K P L Q T C G S P L Y A S P E I V A N G K P Y V G E P V S W S 239
Db      176 L K A N L I L L D A D M N I K I A D P G F S N E F T G N K I D T C G S P Y A A P L F O G K Y D G E P V S W S 235

QY      240 L G V L L Y L I V A G T M P F D Q D H K T L A K O I S N G A Y R E P P K P S D A C - G L I R W L M V N P T R A T L 298
Db      236 L G V I L Y L V G S L P F D Q N L K E L E R V L R G K Y R I P F Y M S I D C E N L L K K F L I L A N S K R G T L 295

QY      299 E D V A S H M W V M V K G Y T T G V G E O A L R E G G H P S G D F P R A S M A D M L R R S S P L L E N G A K V C S F F 358
Db      296 E Q I M K D R M N V N G H - - - E D D E L R P Y V E P L P D Y K P - - - R T E I L M V S M G - - - Y T 338

QY      359 K O H P P G G S T V P G L E R O H S L K K S K E N D M A - - - - - O N L O G D - - - - - P A E D - T S S 401
Db      339 R E E L - - - - - O D S L V G R Y N E V M A I Y L L G Y K S S L E B D T T L L K P R P A D I L N S 386

QY      402 R P G K S I L K L P K G I - - - - - L K K S T S G S G E - - - - - V O E D P O E L R P V 436
Db      387 S A S P S H K V O R S V A N P K O R R F S D O A G P A I P T S N S Y S K T G S N N A E N K R P E D R E S G R K A 446

QY      437 P D T - - - - - R G Q V P V A N S L P R K G I L K K S Q R E S G Y T S S P E S I S G E L L A S D V F V S G D E V E Q 493
Db      447 S S T A K V A S A P L P - - - - - G L E R K K - - - - - T T P P T S N S V L S T S T N - - - - - R S R 483

```

QY 494 KSPQAGLLHHRKGLTK---NGKSRFALGTPPTSGSLDOLASSHPAA-RRSRPSGA 549

Db 484 NSP-----LLRPAISLGGASIQNGKDS-TAPQ-RVPVAPSPSHNITSSSGGAPDRINPFRGV 536

QY 550 VSEDSITLS---ESFPDLDPERL-PEEP 574

Db 537 SSNSTHAGQLQVRDQOQLPFGVTSAPF 565

[illegible]

Qy 546 PSGAVSEDSILSS--ESFDQDLPERL-PETP 574
Db 500 PRGVSSRSTFHAGQLRQVRDQONTLYGVTPASP 532

Search completed: May 11, 2005, 14:26:06
Job time : 46.2557 secs

THIS PAGE BLANK (USP 11)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:09:18 ; Search time 165.602 Seconds

(without alignments)
1473.690 Million cell updates/sec

Title: US-09-980-464-11

Sequence: 1 MESVALLQRPSCAPASASALA.....DCQEVTAAYKALGICSKLS 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_16dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	4	AAB50056 Murine Ly
2	3293	100.0	652	8	ABO84756 Murine ca
3	3068.5	93.2	630	5	AAE19885 Rat SNF1/
4	2824.5	85.8	628	4	AAAG9360 Human pol
5	2824.5	85.8	628	4	ABUS3319 Human cel
6	2824.5	85.8	628	5	ABP69116 Human pol
7	2824.5	85.8	628	5	AAU79652 Human pro
8	2824.5	85.8	628	7	ADP76965 Novel hum
9	2824.5	85.8	628	8	ADL30886 Human pro
10	2824.5	85.8	628	8	ADL25362 Human SNA
11	2824.5	85.8	628	8	ADL14161 Novel hum
12	2824.5	85.8	628	8	ADO20172 Human PRO
13	2824.5	85.8	628	8	ABO84757 Human can
14	2824.5	85.8	628	8	ADJ96620 Human cal
15	2800	85.0	629	4	AAAB5632 Novel pro
16	2800	85.0	629	8	ADJ29239 Human MAR
17	2720.5	82.6	594	5	AAE16266 Human kin
18	2590	78.7	611	6	ABP96085 Human pro
19	2416.5	73.4	534	4	AAAB71959 Human TGF
20	1781.5	54.1	406	5	ABP51382 Human MDD
21	1781.5	54.1	406	5	ABP51482 Human MDD
22	1699.5	51.6	661	4	AAE07847 Human pro
23	1699.5	51.6	661	4	AAE07846 Human pro
24	1699.5	51.6	661	7	ADCE38421 Human pro
25	1699.5	51.6	661	8	ADJ75331 Marker ge

26	1699.5	51.6	661	8	ADL25353 Human ARK
27	1699.5	51.6	661	8	ADQ19734 Human sof
28	1698.5	51.6	660	7	ADN95766 Human BEC
29	1301.5	39.5	530	4	AAE00668 Human pro
30	848	25.8	434	4	ABBE7451 Drosophi1
31	748	22.7	752	4	AAAG3956 Human pol
32	748	22.7	752	4	AAE11782 Human kin
33	748	22.7	752	5	ABBO4433 Human neu
34	748	22.7	752	6	ABG73794 Human MAR
35	748	22.7	752	8	ADL32125 Human pro
36	744	22.6	688	7	ABBO4434 Human neu
37	744	22.6	689	7	ADP74129 Human nov
38	744	22.6	752	8	ADJ96622 Human cal
39	743.5	22.6	769	5	ABP62966 Human pol
40	739	22.4	688	6	AAE33554 Human mic
41	739	22.4	688	6	AAE33555 Human mic
42	739	22.4	688	7	ADG91726 Human mic
43	739	22.4	688	7	ADG91727 Human mic
44	739	22.4	688	8	ADM82131 Tumour-as
45	736	22.4	724	5	ABBO4431 Murine ne

ALIGNMENTS

RESULT 1	ID	AB50056 standard; protein; 631 AA.
XX	AAAB50056;	
XX	19-MAR-2001 (first entry)	
XX	Murine Lymph node Stromal cell kinase 1.	
XX	Murine; Lymph node Stromal cell kinase; MUSK-1; autoimmune disorder;	
XX	wound healing; periodontal disease; inflammatory disease; tumour;	
XX	infection; allergy.	
XX	Mus musculue.	
XX	WO200073468-A1.	
XX	07-DEC-2000.	
XX	26-MAY-2000; 2000WO-US014696.	
XX	28-MAY-1999; 99US-0136781P.	
XX	(IMMV) IMMUNEX CORP.	
XX	Bird TA, Virca GD, Martin U, Anderson DM;	
XX	WPI; 2001-061546/07.	
XX	N-PSDB; AAC90433.	
PT	Novel murine and human kinase nucleic acids useful for treating	
PT	infections, tumors, allergies, autoimmune diseases, and	
PT	for stimulating or suppressing immune responses.	
PS	Claim 10; Page 94-96; 106pp; English.	
CC	The present sequence is Murine Lymph node Stromal cell kinase 1 (MUSK-1).	
CC	This protein is useful for treating a variety of disorders listed in the	
CC	disclosure of the specification, including autoimmune disorders, allergic	
CC	reactions, myeloid or lymphoid cell deficiencies, wound healing, and	
CC	tissue repair and replacement, burns, incisions and ulcers, periodontal	
CC	disease, inflammatory diseases, tumours and bacterial, viral or fungal	
CC	infection	
XX	Sequence 631 AA;	
XX	Query Match 100.0%; Score 3293; DB 4; Length 631;	

Best Local Similarity 100.0%; Pred. No. 1e-280;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQASALASASARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
PT |||||
DB 1 MESVALLQRPSPQASALASASARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
QY 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHIIAH 120
DB 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHIIAH 120
QY 121 EVFENSKIYIWEYASRGDLVDYISERPRLSERDARHPROIVSALHYCHONGIYHRDL 180
DB 121 EVFENSKIYIWEYASRGDLVDYISERPRLSERDARHPROIVSALHYCHONGIYHRDL 180
QY 181 KLENILLDANGNIKIADFGISNLVYHKGKFLQTFGSPFLVASPEIVNGKPYVGEVDSWSL 240
DB 181 KLENILLDANGNIKIADFGISNLVYHKGKFLQTFGSPFLVASPEIVNGKPYVGEVDSWSL 240
QY 241 GVLVYLVHGTMPDQDHTLVKQISNGAYRREPSPKSDACGLIRWLLMVPTRRATLED 300
DB 241 GVLVYLVHGTMPDQDHTLVKQISNGAYRREPSPKSDACGLIRWLLMVPTRRATLED 300
QY 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
DB 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
QY 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
DB 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
QY 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIKKSRQSSGYSSPEPESGELLD 480
DB 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIKKSRQSSGYSSPEPESGELLD 480
QY 481 ASDVFPVSGDPVEQSPQASGLLHRKGIKLNKGFSTALEGTPSTFGSLDQLASSHPA 540
DB 481 ASDVFPVSGDPVEQSPQASGLLHRKGIKLNKGFSTALEGTPSTFGSLDQLASSHPA 540
QY 541 ARPSRPSGAVSEDSIISSESFDQDLPERLPETPLRGCVSVNMLRGLEQPPSEGLKRWQ 600
DB 541 ARPSRPSGAVSEDSIISSESFDQDLPERLPETPLRGCVSVNMLRGLEQPPSEGLKRWQ 600
QY 601 ESLGDSCSFSLTDCQEVTAAYRQALGICSKLS 631
DB 601 ESLGDSCSFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 2
ABO84756
ID ABO84756 standard; protein; 652 AA.
XX
AC ABO84756;
XX
DT 18-NOV-2004 (first entry)
XX
DB Murine cancer-associated protein (CAP) M07-004.
XX
KW Mouse; cancer-associated protein; CAP; cancer; cytostatic.
XX
OS Mus musculus.
XX
PN MO2004068146-A2.
XX
PD 15-JUL-2004.
XX
PF 15-DEC-2003; 2003MO-US040081.
XX
PR 17-DEC-2002; 2002US-00322281.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;

XX WIPI, 2004-499109/47.
DR N-PSDB; ABD33083.
XX
PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Disclosure, SEQ ID NO 23; 162pp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CAP of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 652 AA;
XX
Query Match 100.0%; Score 3293; DB 8; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.e-280;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQASALASASARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
DB 22 MESVALLQRPSPQASALASASARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 81
QY 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHIIAH 120
DB 82 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHIIAH 141
QY 121 EVFENSKIYIWEYASRGDLVDYISERPRLSERDARHPROIVSALHYCHONGIYHRDL 180
DB 142 EVFENSKIYIWEYASRGDLVDYISERPRLSERDARHPROIVSALHYCHONGIYHRDL 201
QY 181 KLENILLDANGNIKIADFGISNLVYHKGKFLQTFGSPFLVASPEIVNGKPYVGEVDSWSL 240
DB 202 KLENILLDANGNIKIADFGISNLVYHKGKFLQTFGSPFLVASPEIVNGKPYVGEVDSWSL 261
QY 241 GVLVYLVHGTMPDQDHTLVKQISNGAYRREPSPKSDACGLIRWLLMVPTRRATLED 300
DB 262 GVLVYLVHGTMPDQDHTLVKQISNGAYRREPSPKSDACGLIRWLLMVPTRRATLED 321
QY 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
DB 322 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 381
QY 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
DB 382 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 441
QY 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIKKSRQSSGYSSPEPESGELLD 480
DB 442 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIKKSRQSSGYSSPEPESGELLD 501
QY 481 ASDVFPVSGDPVEQSPQASGLLHRKGIKLNKGFSTALEGTPSTFGSLDQLASSHPA 540
DB 502 ASDVFPVSGDPVEQSPQASGLLHRKGIKLNKGFSTALEGTPSTFGSLDQLASSHPA 561
QY 541 ARPSRPSGAVSEDSIISSESFDQDLPERLPETPLRGCVSVNMLRGLEQPPSEGLKRWQ 600
DB 562 ARPSRPSGAVSEDSIISSESFDQDLPERLPETPLRGCVSVNMLRGLEQPPSEGLKRWQ 621
QY 601 ESLGDSCSFSLTDCQEVTAAYRQALGICSKLS 631


```
XX Human polypeptide, SEQ ID NO: 2919.
DE
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping.
KW
XX Homo sapiens.
OS
XX EP130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-00114089.
PF
XX
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
PI
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94280.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 2919; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a polypeptide encoded by a full length
XX human cDNA of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in CD-ROM
XX format directly from EPO
XX
XX Sequence 628 AA;
SQ
Query Match 85.8%; Score 2824.5; DB 4; length 628;
Best Local Similarity 85.7%; Pred. No. 2e-239;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLQRPSPQSPSALASASARPLADGLIKSPKPLMKQAVKCHHKNTRHRYEFL 60
DB 1 MESLVFARRSGPTPS----AAELARPLAEGLIKSPKPLMKQAVKCHHKNTRHRYEFL 56
QY 61 ETLGKGTGKVKKARSSGRLVAIKIRKOKIKDEBDLHIREIFIMSSLNPHI1AH 120
DB 57 ETLGKGTGKVKKARSSGRLVAIKIRKOKIKDEBDLHIREIFIMSSLNPHI1AH 116
QY 121 EVFENSCKIVIVMEYASRGDLVYISERPLSRDARHFRQIVASLHYCHQNGIYHRDL 180
DB 117 EVFENSCKIVIVMEYASRGDLVYISERQLSEREAHFRQIVASLHYCHQNGIYHRDL 176
QY 181 KLENILLDANGNINKIDFGISNLYHKGKPLQTCGSPLYASPIIVNGKPYGPEVDWSL 240
DB 177 KLENILLDANGNINKIDFGISNLYHKGKPLQTCGSPLYASPIIVNGKPYGPEVDWSL 236
QY 241 GVLVLYLVHGTMPFDGQDHKTLVKOISNGAYRPPSPDAICGILRWLWMPRRRATLTD 300
DB 237 GVLVLYLVHGTMPFDGQDHKTLVKOISNGAYRPPSPDAICGILRWLWMPRRRATLTD 296
QY 301 VASHMVMVNGYTTGVGEORALREGHPSGDFGRASADWLRRSSRPLLENGATVCSFFKQ 360
DB 297 VASHMVMVNGYTTGVGEORALREGHPSGDFGRASADWLRRSSRPLLENGATVCSFFKQ 356
```

```
QY 361 HVPGGSTVPGLEGRQSLKSKRENDMAQNLQDPADETSRRPGKSLKPKGILKKSS 420
DB 357 HAPGGSTVPGLEGRQSLKSKRENDMAQNLQDPADETSRRPGKSLKPKGILKKSS 416
QY 421 TSSGEVQEDPQELRPVDPTRGQVPVAVSLPRKGIIKKSRQSRGSSYSSPEPSGELD 480
DB 417 ASAEVQEDPPELSPFIPASGQMAP--LIPKGIILKKPRQSRGSSYSSPEPSGELD 473
QY 481 ASDVFSVGDPEVQKSPQASGLLHRRGIIKLKNGKFSRTALEGTPPTSGSLDOLASSHPA 540
DB 474 AGDVFSVGDPEVQKSPQASGLLHRRGIIKLKNGKFSQTLALAPPTFGSLDELAPRPL 533
QY 541 ARPSRPSGAVSEDSIISSESPDLDLPERLPEPLPLKGVSVNDLRLGLEPPSGS----LK 536
DB 534 ARASRPSGAVSEDSIISSESPDLDLPERLPEPLPLKGVSVNDLRLGLEPPSGSGLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYROALGICSLT 631
DB 594 RWRQDPLGDSCFSLTDCQEVTAAYROALRVCSKLT 628
```

RESULT 5

ABUS3319
ID ABUS3319 standard; protein; 628 AA.

AC ABUS3319;

DT 14-APR-2003 (first entry)

DE Human cell cycle-associated protein from DKFzphes3_7j3.

KW Human; gene therapy; vaccine; disease treatment; detection.

OS Homo sapiens.

PN WO200112659-A2.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-1B001496.

PR 18-AUG-1999; 99US-0149499P.

PR 28-SEP-1999; 99US-0156503P.

XX (GENU-) GERMAN HUMAN GENOME PROJECT.

PA Wiemann S;

PI WPI; 2001-327840/34.

DR N-PSDB; ABX71420.

PT Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies.

PS Claim 21; Page 943; 1095pp; English.

XX This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a polypeptide described in the disclosure of the invention

XX Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 4; length 628;
Best Local Similarity 85.7%; Pred. No. 2e-239;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

DB 474 AGDVFSGDPKQKPPQASGLLHRRKGIILKNGKFSQTALFLAAPTFFGSLDELAPRPL 533
QY 541 ARSRPSGAVSSESISSSEFPQDLPERLPETPLRGCVSVNLRGLRQPPSG---LK 596
DB 534 ARSRPSGAVSSESISSSEFPQDLPERLPETPLRGCVSVNLRGLRQPPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 7
AAU79652
AAU79652 standard; protein; 628 AA.
XX
AC AAU79652;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human protein kinase 3700.
XX
KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
KW cell proliferation disorder; cell differentiation disorder; carcinoma;
KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;
KW cytoskeletal; antiatherosclerotic; enzyme.
XX
OS Homo sapiens.
XX
PN MO20024921-A2.
XX
PD 28-MAR-2002.
XX
PF 25-SEP-2001; 2001MO-US030115.
XX
PR 25-SEP-2000; 2000US-0234922P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Galvin KM;
XX
DR MPI; 2002-352007/38.
XX
DR N-PSDB; ABK14000.
XX
PT Use of modulators of activity of 3700 protein for making medicament for
PT e.g., modulating protein phosphorylation or cell signalling, or for
PT treating or preventing cellular proliferative and/or differentiative
PT disorders.
XX
PS Claim 19; Fig 1; 115pp; English.

CC The present invention relates to the isolation of a novel human protein
CC kinase designated 3700, and the polynucleotide sequence encoding it. The
CC invention also describes the use of a modulator of the activity of
CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
CC composition for modulating the ability of a cell to phosphorylate an
CC amino acid residue of a substrate protein. Modulators of protein kinase
CC 3700 activity are useful for modulating protein phosphorylation, cell
CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
CC angiogenesis, tissue repair, tissue regeneration, establishment or
CC progression of atherosclerosis, and signalling across the blood-brain
CC barrier. The polynucleotide and polypeptide molecules for protein kinase
CC 3700 may be used as diagnostic targets and therapeutic agents for
CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
CC curing PK-related disorders and cellular proliferative and/or
CC differentiative disorders (e.g. haematopoietic neoplastic disorders,
CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
CC polynucleotide sequence can be used to express protein kinase 3700, to
CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for
CC tissue typing, in forensic biology, and as surrogate markers. The present
CC sequence represents human protein kinase 3700

XX
SQ Sequence 628 AA;
Query Match 85.8%; Score 2824.5; DB 5; Length 628;
Best Local Similarity 85.7%; Pred. No. 2e-239;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLQRPQAPSAALASASARLADGLISPPPLMKKQAVKXHHKHNIRHYEFL 60
DB 1 MESLVAPARRSGPTPS---AAILARPLABGLIKSPKPKKKQAVKXHHKHNIRHYEFL 56
QY 61 ETLGKGTGKVKKARSSGRVLVAIKSIRKDKIADBDLHIREIRIMSLNPHIAT 120
DB 57 ETLGKGTGKVKKARSSGRVLVAIKSIRKDKIADBDLHIREIRIMSLNPHIAT 116
QY 121 EVFENSSKIIVIMEVASRGDLYDISEPRPLSERDARHFRQIVSAHYCHONGIYRDL 180
DB 117 EVFENSSKIIVIMEVASRGDLYDISEPRPLSERDARHFRQIVSAHYCHONGIYRDL 176
QY 181 KLENILLDANGNIKIDFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVDSWSL 240
DB 177 KLENILLDANGNIKIDFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVDSWSL 236
QY 241 GVLITYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGILRWLWNPTRATLBD 300
DB 237 GVLITYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGILRWLWNPTRATLBD 296
QY 301 VASHMWMVNGYTTGVGEQALREGGHPSGDFGRASNAWMLRRSSRPLTENGAVCSEFFQ 360
DB 297 VASHMWMVNGYATRVGEQALREGGHPSGDFGRASNAWMLRRSSRPLTENGAVCSEFFQ 356
QY 361 HYPGGGSTYVGLERQSLKSKRENDMAONLQDPAEDTSRRPKGSLKLPKILKKKS 420
DB 357 HAPGGGSTYVGLERQSLKSKRENDMAQSLSDTADTADHRRGKSNLKLPKILKKKS 416
QY 421 TSSGEVQEDPOELRPVDPTRGQVPVPAVSLPRKGIKKRSORESGYSSPPESEGLD 480
DB 417 ASHGVQEDPPELSIPSPASPGQAP--LPPKGIKKRQKRSBGYSPPPESEGLD 473
QY 481 ASDVFSGDPVEOKSPQASGLLHRRKGIILKNGKFSQTALFLAAPTFFGSLDELASHPA 540
DB 474 AGDVFSGDPKQKPPQASGLLHRRKGIILKNGKFSQTALFLAAPTFFGSLDELAPRPL 533
QY 541 ARSRPSGAVSSESISSSEFPQDLPERLPETPLRGCVSVNLRGLRQPPSG---LK 596
DB 534 ARSRPSGAVSSESISSSEFPQDLPERLPETPLRGCVSVNLRGLRQPPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 8
ADF76965
ADF76965 standard; protein; 628 AA.
XX
AC ADF76965;
XX
DT 26-FEB-2004 (first entry)
XX
DE Novel human secreted and transmembrane protein SeqID 640.
XX
KW human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytoskeletal; chondrocyte; tumour.
XX
OS Homo sapiens.
XX
PN MO2003072035-A2.
XX
PD 04-SEP-2003.

Db 237 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKPSDACGLIRMLVMNPTTRATLED 236
QY 301 VASHMWNMGYYTTGVBGEQALREGGHPGSGFGRASNADWLRRSGRPLLENGAKVCSFFKQ 360
Db 297 VASHMWNMGYYATRVGEQAPHEGHPGSGDSARASNADWLRRSSRPLLENGAKVCSFFKQ 356
QY 361 HVGGSSTVGLERHSHLSKSRKENDMAONLQDPAEDTSRRPKSSLKLPKGIILKKKS 420
Db 357 HAPGGSTTPGLRQHSLSKSRKENDMAQSLHSDTADTAAHRFGKSNLKPKGILKKKVS 416
QY 421 TSSGEVQEDPQELRPVDPTRFGQVPVAVSLPRKGIILKKSRQSRGYSSEPESEGEILD 480
Db 417 ASAEVQEDPEPLSPIPASPGQAP---LLPKGILKKRQSRGYSSEPESEGEILD 473
QY 481 ASDVFGDPVPEQKSPQASGLLHRKGIILKNGKFSRTALEGTTPTFGSLDQLASHPA 540
Db 474 AGVDFVSGDPKPKPPQASGLLHRKGIILKNGKFSQTLAELAPPTFGSLDELAPRPL 533
QY 541 ARSPRPSGAVSEDSILSSSEFPDQDLPRLPETPLRGCVSVNLTGLEBPSPSGCLR 596
Db 534 ARSPRPSGAVSEDSILSSSEFPDQDLPRLPETPLRGCVSVNLTGLEBPSPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSCFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 12

ADO20172
ID ADO20172 standard; protein; 628 AA.

AC ADO20172;

DT 12-AUG-2004 (first entry)

DE Human PRO polypeptide #540.

Human; PRO; immune related disorder; systemic lupus erythematosus;
rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
diabetes mellitus; renal disease; demyelinating disease;
central nervous system; peripheral nervous system;
demyelinating polyneuropathy; Guillain-Barre syndrome;
chronic inflammatory demyelinating polyneuropathy.

OS Homo sapiens.

PN MO2004043361-AZ.

PD 27-MAY-2004.

PF 06-NOV-2003; 2003WO-US035268.

PR 08-NOV-2002; 2002US-0425235P.

PA (GENTR) GENENTECH INC.

PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM,

PI Wood WI, Wu TD;

DR WPI; 2004-420067/39.

DR N-PSDB; ADO20171.

Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO68388 useful for
treating an immune related disorder such as systemic lupus erythematosus,
rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
spondyloarthritis.

Claim 7; SEQ ID NO 1080; 1731pp; English.

The invention relates to human PRO polypeptides and the polynucleotides
encoding them. The polypeptides and polynucleotides are useful for

treating and diagnosing immune related disorders in mammals. The immune
related disorders include systemic lupus erythematosus, rheumatoid
arthritis, osteoarthritis, juvenile chronic arthritis, systemic
sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
mellitus, immune-mediated renal disease, demyelinating diseases of the
central or peripheral nervous system, demyelinating polyneuropathy,
Guillain-Barre syndrome and chronic inflammatory demyelinating
polyneuropathy. This sequence represents a human PRO polypeptide of the
invention.

SQ Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;

Best Local Similarity 85.7%; Pred. No. 26-239;

Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLORPSQAPASALASESARPLAAGLTKSPKPLMKQAVRHHKHLRHYEPL 60
Db 1 MESLVFARRSGTPPS---AAELARPLAEGLTKSPKPLMKQAVRHHKHLRHYEPL 56
QY 61 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEODLHIREIEIMSLNHPHIIAT 120
Db 57 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEODLHIREIEIMSLNHPHIIAT 116
QY 121 EVFENSSKIYIWEIARSDLYDYISERPLSEBDARHFRQIVSAIHYCHONGIYHRL 180
Db 117 EVFENSSKIYIWEIARSDLYDYISERQULSEREARHFRQIVSAIHYCHONGIYHRL 176
QY 181 KLENTLLDANGNIKADPGLSNLHYHKGKFLQFFGCSPLVASEIYNGKRYVPEVDSWL 240
Db 177 KLENTLLDANGNIKADPGLSNLHYHKGKFLQFFGCSPLVASEIYNGKRYTPEVDSWL 236
QY 241 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKPSDACGLIRMLVMNPTTRATLED 300
Db 237 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKPSDACGLIRMLVMNPTTRATLED 296
QY 301 VASHMWNMGYYTTGVBGEQALREGGHPGSGFGRASNADWLRRSGRPLLENGAKVCSFFKQ 360
Db 297 VASHMWNMGYYATRVGEQAPHEGHPGSGDSARASNADWLRRSSRPLLENGAKVCSFFKQ 356
QY 361 HVGGSSTVGLERHSHLSKSRKENDMAONLQDPAEDTSRRPKSSLKLPKGIILKKKS 420
Db 357 HAPGGSTTPGLRQHSLSKSRKENDMAQSLHSDTADTAAHRFGKSNLKPKGILKKKVS 416
QY 421 TSSGEVQEDPQELRPVDPTRFGQVPVAVSLPRKGIILKKSRQSRGYSSEPESEGEILD 480
Db 417 ASAEVQEDPEPLSPIPASPGQAP---LLPKGILKKRQSRGYSSEPESEGEILD 473
QY 481 ASDVFGDPVPEQKSPQASGLLHRKGIILKNGKFSRTALEGTTPTFGSLDQLASHPA 540
Db 474 AGVDFVSGDPKPKPPQASGLLHRKGIILKNGKFSQTLAELAPPTFGSLDELAPRPL 533
QY 541 ARSPRPSGAVSEDSILSSSEFPDQDLPRLPETPLRGCVSVNLTGLEBPSPSGCLR 596
Db 534 ARSPRPSGAVSEDSILSSSEFPDQDLPRLPETPLRGCVSVNLTGLEBPSPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSCFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 13

ABO84757
ID ABO84757 standard; protein; 628 AA.

AC ABO84757;

DT 18-NOV-2004 (first entry)

DE Human cancer-associated protein (CAP) HP07-004.

Human; cancer-associated protein; CAP; cancer; cytostatic.

XX Homo sapiens.
 XX MO2004058146-A2.
 XX
 PD 15-JUL-2004.
 XX
 PF 15-DEC-2003; 2003MO-US040081.
 XX
 PR 17-DEC-2002; 2002US-00322281.
 XX
 PA (SAGR-) SAGES DISCOVERY INC.
 PI Morris DW, Malandro MS;
 PI WPI; 2004-499109/47.
 DR N-PSDB; ABDJ3085.
 XX
 PT Novel human cancer associated protein encoded within open reading frame
 of cancer associated gene, useful as targets for diagnosing cancer.
 XX
 PS Claim 18; SEQ ID NO 26; 182bp; English.
 XX
 CC The invention relates to cancer-associated proteins (CAP) and the cancer-
 associated (CA) nucleic acids encoding them. The invention also relates
 to a method for treating cancers involving administering to a patient an
 inhibitor of CAP, and a method of screening for anticancer activity in a
 potential drug involving providing a cell that expresses a CA gene,
 contacting a tissue sample derived from a cancer cell with an anticancer
 drug candidate and monitoring the effect of the anticancer drug candidate
 on expression of the CA gene. The CAP proteins are useful for detecting
 cancer associated with expression of a CAP protein in a test cell sample
 and for screening for a bioactive agent capable of modulating the
 activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 cancer, involving determining the expression of a CA nucleic acid in a
 tissue. This sequence represents a human CAP of the invention. Note: The
 sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 628 AA;
 Query Match 85.8%; Score 2824.5; DB 8; Length 628;
 Best Local Similarity 85.7%; Pred. No. 26-239;
 Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
 QY 1 MESVALLQRPSPQASALASASAPLADGLISPPRLMKQAVKHHKHNIRHYEFL 60
 DB 1 MESLVFARRSGPTPS---AAELARPLAEGLIKSPKLMKQAVKHHKHNIRHYEFL 56
 QY 61 ETLGKGTGKVKKARRSGRLVAIKIRKDKIDBDLHIREIEMSLNPHITAH 120
 DB 57 ETLGKGTGKVKKARRSGRLVAIKIRKDKIDBDLHIREIEMSLNPHITAH 116
 QY 121 EYFENSSTIVIMEVASRGLDYISERPRISRDAHFPRQIVASLHYCHQGIYHRDL 180
 DB 117 EYFENSSTIVIMEVASRGLDYISERQLSREARHFRQIVASVAVHYCHQRIYHRDL 176
 QY 181 KLENIILLDANGNIKIDPGLISNLVHKGKFLQTFCSPLVASPIVNGKPYGPEVDWSL 240
 DB 177 KLENIILLDANGNIKIDPGLISNLVHKGKFLQTFCSPLVASPIVNGKPYGPEVDWSL 236
 QY 241 GVLVLYLVHGTMPFDGDHKTIVKOISNGAYRPPPSDGLIRWLMVNPFRRAITLED 300
 DB 237 GVLVLYLVHGTMPFDGDHKTIVKOISNGAYRPPPSDGLIRWLMVNPFRRAITLED 296
 QY 301 VASHMWVNWGYTTGVGEQEARLREGGHPSGDFGRASMAWMLRRSSRPLLENGATVCSFFKQ 360
 DB 297 VASHMWVNWGYTTGVGEQEARLREGGHPSGDFGRASMAWMLRRSSRPLLENGATVCSFFKQ 356
 QY 361 HYPGGSGTVPGLERQSLKSKRENDAQNLQDPAEDTSSRPKGSLLKPKIILKKSS 420
 DB 357 HAPGGSGTVPGLERQSLKSKRENDAQNLQDPAEDTSSRPKGSLLKPKIILKKSS 416

QY 421 TSSGGEVQEDPQELRPVDTPPQGVPAVSLPKKGIILKKSRORESGYSSPEPSGGLLD 480
 DB 417 ASAEQVQEDPPELSPIPASPGQAP---LLPKGGIILKKSRORESGYSSPEPSGGLLD 473
 QY 481 ASDPVFSGDPEKESQASGLLHRRGIIILKKNKFSRTALEGTTPTFGSLDOLASHHPA 540
 DB 474 AGDVFVSGDPEKQKPPQASGLLHRRGIIILKKNKFSQTALSLAAPTTPGSLDELAPRPL 533
 QY 541 ARPSRPGAVSEDSIISSESPDLDPERLPEPTLRGCVSVNDLARGLEOPPSGEG---LK 536
 DB 534 ARASRPGAVSEDSIISSESPDLDPERLPEPTLRGCVSVNDLARGLEOPPSGEGCLR 533
 QY 597 RWMQESLGDSCPSGLTDCQEVTAAYRQALGICSKLSS 631
 DB 594 RWRQDPLGDSCFSLTDCQEVTAAYRQALRVCSKLT 628
 RESULT 14
 ADJ96620
 ID ADJ96620 standard; protein; 672 AA.
 XX
 AC ADJ96620;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human calcium/calmodulin-dependent protein kinase NuaK2 protein Seqid 77.
 XX
 KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
 KW PK; ST; gene therapy; cancer; immune-related disease;
 KW cardiovascular disease; brain; neuronal associated disease; metabolic;
 KW inflammatory disorder; cytosolic; neuroprotective; immunomodulator;
 KW antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;
 KW NuaK2.
 XX
 OS Homo sapiens.
 OS 72.
 PN MO2004006838-A2.
 PD 22-JUN-2004.
 XX
 PF 15-JUL-2003; 2003MO-US021730.
 XX
 PR 15-JUL-2002; 2002US-0395632P.
 XX
 PA (SUGR-) SUGEN INC.
 PI Whyte D, Manning G, Caenepeel S;
 PI WPI; 2004-122753/12.
 DR N-PSDB; ADJ96554.
 XX
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for
 preparing a composition for treating diseases or disorders, e.g., cancer,
 or neurological, immunological or inflammatory disorders.
 XX
 PS Claim 1; SEQ ID NO 77; 366bp; English.
 XX
 CC This invention relates to a novel isolated, enriched or purified nucleic
 acid molecule that encodes a kinase polypeptide. Specifically, it relates
 to human tyrosine and serine/threonine protein kinases (PK's and STK's),
 as well as protein kinase-like enzymes. The present invention describes
 screening methods to identify agonists, antagonists and antibodies that
 can be used to modulate the activity or function of the mammalian kinase
 enzymes. As such, these compositions can be used for gene therapy
 purposes to treat diseases or disorders including cancer, immune-related
 diseases, cardiovascular disease, brain or neuronal associated disease,
 metabolic and inflammatory disorders. Accordingly, they exhibit
 cytostatic, neuroprotective, immunomodulator and antiinflammatory
 activities. This polypeptide sequence is a human kinase protein sequence
 of the invention.

Qy 480 DASDVFSGDVYBOKSPQASGILLHRKGIKLNKFSRTALBGTTPSTFGSLDQIASSHP 539
Db 474 DAGDVFSGDVYBOKSPQASGILLHRKGIKLNKFSQTALBAPPTFGSLDEIAPPRP 533
Qy 540 AARPSPSGAVSGSDSTLSSSEFPQDLPLPRLPETPLRGCVSVNLTGLBOPPEG----L 595
Db 534 LARASRPSGAVSGSDSTLSSSEFPQDLPLPRLPETPLRGCVSVNLTGLBOPPEGSGCL 593
Qy 596 KRWQESLGDSCFSLTDCOEVTAAVRAALGICSKLS 631
Db 594 RRRQDPPLGDSCTSLTDCOEVTATYRQALRVCSKLT 629

Search completed: May 11, 2005, 14:20:07
Job time : 171.602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 11, 2005, 14:24:11 ; Search time 129.912 Seconds

(Without alignments)
1622.537 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293
Sequence: 1 MESVALLQRPSPQAPSASALA.....DCQEVTAAYRQALGICSKLS 631Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	14	US-10-355-975-11 Sequence 11, Appl
2	3186	96.8	640	16	US-10-322-281-23 Sequence 23, Appl
3	3068.5	93.2	630	16	US-10-343-514-41 Sequence 41, Appl
4	2824.5	85.8	628	9	US-09-963-159-2 Sequence 2, Appl
5	2824.5	85.8	628	15	US-10-423-543-44 Sequence 44, Appl
6	2824.5	85.8	628	16	US-10-618-941-77 Sequence 77, Appl
7	2720.5	82.4	594	15	US-10-311-034-12 Sequence 12, Appl
8	2712.5	82.6	616	16	US-10-322-281-26 Sequence 26, Appl
9	1781.5	54.1	406	16	US-10-363-829-404 Sequence 404, App
10	1781.5	54.1	406	16	US-10-363-829-504 Sequence 504, App
11	1699.5	51.6	661	9	US-09-780-949-2 Sequence 2, Appl
12	1699.5	51.6	661	9	US-09-780-949-6 Sequence 2, Appl
13	1699.5	51.6	661	14	US-10-354-358-82 Sequence 82, Appl

14	1325	40.2	251	16	US-10-343-514-103	Sequence 103, App
15	1301.5	39.5	530	9	US-09-836-392-20	Sequence 20, Appl
16	1296	-39.4	251	16	US-10-343-514-50	Sequence 50, Appl
17	1276	38.7	251	16	US-10-343-514-101	Sequence 101, App
18	1111.5	33.8	252	16	US-10-343-514-102	Sequence 102, App
19	785.5	23.9	246	10	US-09-898-837A-29	Sequence 29, Appl
20	748	22.7	703	15	US-10-016-248-70	Sequence 70, Appl
21	748	22.7	752	9	US-09-835-081-2	Sequence 2, Appl
22	748	22.7	752	15	US-10-258-106-16	Sequence 16, Appl
23	748	22.7	752	15	US-10-276-645-7	Sequence 7, Appl
24	748	22.7	825	15	US-10-425-114-54516	Sequence 54516, A
25	744	22.6	639	15	US-10-016-248-71	Sequence 71, Appl
26	744	22.6	639	15	US-10-016-248-72	Sequence 72, Appl
27	744	22.6	688	15	US-10-276-645-8	Sequence 8, Appl
28	744	22.6	752	16	US-10-618-941-79	Sequence 79, Appl
29	743.5	22.6	769	15	US-10-363-616-403	Sequence 403, Appl
30	739	22.4	639	15	US-10-016-248-73	Sequence 73, Appl
31	739	22.4	688	14	US-10-161-565-28	Sequence 28, Appl
32	739	22.4	688	14	US-10-161-565-29	Sequence 29, Appl
33	736	22.4	724	15	US-10-276-645-5	Sequence 5, Appl
34	732.5	22.2	660	15	US-10-276-645-6	Sequence 6, Appl
35	732.5	21.4	729	14	US-10-142-356-11	Sequence 11, Appl
36	703.5	21.4	729	14	US-10-195-101-33	Sequence 33, Appl
37	703.5	21.4	729	14	US-10-161-565-26	Sequence 26, Appl
38	701	21.3	140	16	US-10-343-514-47	Sequence 47, Appl
39	701	21.3	744	9	US-09-835-081-4	Sequence 4, Appl
40	700.5	21.3	1038	15	US-10-424-599-274878	Sequence 274878, A
41	698.5	21.2	624	15	US-10-425-114-70100	Sequence 70100, A
42	698.5	21.2	744	9	US-09-919-585-3	Sequence 3, Appl
43	698	21.2	698	15	US-10-016-248-22	Sequence 22, Appl
44	696.5	21.2	508	15	US-10-016-248-24	Sequence 24, Appl
45	694.5	21.1	667	15	US-10-016-248-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-10-355-975-11
; Sequence 11, Application US/10355975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/10/355,975
; PRIOR FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579,664B
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975-11

Query Match 100.0%; Score 3293; DB 14; Length 631;
Best Local Similarity 100.0%; Pred. No. 1, le-206;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQAPSASALASARPLADGLIKSPKLMKKQAVKXHHKHNLRHYEFL 60
DB 1 MESVALLQRPSPQAPSASALASARPLADGLIKSPKLMKKQAVKXHHKHNLRHYEFL 60
QY 61 ETIGKTYGVKVKARSSGRLVAIKIRKDKIKDODLHIREITMSLNPHITAIH 120
DB 61 ETIGKTYGVKVKARSSGRLVAIKIRKDKIKDODLHIREITMSLNPHITAIH 120

```
QY 121 EVFENSKIYIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIYVRDL 180
| | | | |
DB 121 EVFENSKIYIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIYVRDL 180
QY 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNKPYVGEVDSWSL 240
| | | | |
DB 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNKPYVGEVDSWSL 240
QY 241 GULLYILVHGTMPFDGDHKTIVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
| | | | |
DB 241 GULLYILVHGTMPFDGDHKTIVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
QY 301 VASHMMVNMNGYTTGVGEQELREGGHPSGDPRASMAWMLRRSSRPLLENGAVCSFFKQ 360
| | | | |
DB 301 VASHMMVNMNGYTTGVGEQELREGGHPSGDPRASMAWMLRRSSRPLLENGAVCSFFKQ 360
QY 361 HVPGGGSTVPGELRQSLKKSREKENDMAQNLQGDPAEDTSSRPKSSLKLPKGI LKXKS 420
| | | | |
DB 361 HVPGGGSTVPGELRQSLKKSREKENDMAQNLQGDPAEDTSSRPKSSLKLPKGI LKXKS 420
QY 421 TSSGEVQEDPOELRPVPDTPGQVPVAVSLPRKGI LKKSQRBSGYSSPESPESGELLD 480
| | | | |
DB 421 TSSGEVQEDPOELRPVPDTPGQVPVAVSLPRKGI LKKSQRBSGYSSPESPESGELLD 480
QY 481 ASDVFPVSGDPVEQKSPQASGLLHRRKGI LKNGKFSRTALEGTTPTFSGLDQLASSHPA 540
| | | | |
DB 481 ASDVFPVSGDPVEQKSPQASGLLHRRKGI LKNGKFSRTALEGTTPTFSGLDQLASSHPA 540
QY 541 ARSPRPSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNDLRGLEQPPSEGLKRWMO 600
| | | | |
DB 541 ARSPRPSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNDLRGLEQPPSEGLKRWMO 600
QY 601 ESI GDSCFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 601 ESI GDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 2
US-10-322-281-23
; Sequence 23, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PaateSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-322-281-23

Query Match 96.8%; Score 3186; DB 16; Length 640;
Best Local Similarity 98.3%; Pred. No. 1.1e-199;
Matches 620; Conservative 0; Mismatches 1; Indels 10; Gaps 5;
```

```
DB 192 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLY-DEIVNGKPYVGEVDSWSL 249
| | | | |
QY 241 GULLYILVHGTMPFDGDHKTIVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
| | | | |
DB 250 GULLYILVHGTMPFDGDHKTIVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 309
| | | | |
QY 301 VASHMMVNMNGYTTGVGEQELREGGHPSGDPRASMAWMLRRSSRPLLENGAVCSFFKQ 360
| | | | |
DB 310 VASHMMVNMNGYTTGVGEQELREGGHPSGDPRASMAWMLRRSSRPLLENGAVCSFFKQ 369
QY 361 HVPGGGSTVPGELRQSLKKSREKENDMAQNLQGDPAEDTSSRPKSSLKLPKGI LKXKS 420
| | | | |
DB 370 HVPGGGSTVPGELRQSLKKSREKENDMAQNLQGDPAEDTSSRPKSSLKLPKGI LKXKS 429
QY 421 TSSGEVQEDPOELRPVPDTPGQVPVAVSLPRKGI LKKSQRBSGYSSPESPESGELLD 480
| | | | |
DB 430 TSSGEVQEDPOELRPVPDTPGQVPVAVSLPRKGI LKKSQRBSGYSSPESPESGELLD 489
QY 481 ASDVFPVSGDPVEQKSPQASGLLHRRKGI LKNGKFSRTALEGTTPTFSGLDQLASSHPA 540
| | | | |
DB 490 ASDVFPVSGDPVEQKSPQASGLLHRRKGI LKNGKFSRTALEGTTPTFSGLDQLASSHPA 549
QY 541 ARSPRPSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNDLRGLEQPPSEGLKRWMO 600
| | | | |
DB 550 ARSPRPSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNDLRGLEQPPSEGLKRWMO 609
QY 601 ESI GDSCFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 610 ESI GDSCFSLTDCQEVTAAYRQALGICSKLS 640
```

```
RESULT 3
US-10-343-514-41
; Sequence 41, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBEVRRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 630
; TYPE: PRT
; ORGANISM: RAT
US-10-343-514-41
```

```
Query Match 93.2%; Score 3068.5; DB 16; Length 630;
Best Local Similarity 94.0%; Pred. No. 4.9e-192;
Matches 593; Conservative 12; Mismatches 25; Indels 1; Gaps 1;
```

```
QY 1 MESVALLORPSQAPSAALASAPPLADGLIKSPKPMKQAVXKHHKHLRHRVEFL 60
| | | | |
DB 1 MESVALLHRRGNLAPSAALATESAPPLADRLIKSPKPMKQAVXKHHKHLRHRVEFL 60
QY 61 ETLGKGTGYKVKKARSSSGLVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAIIH 120
| | | | |
DB 61 ETLGKGTGYKVKKARSSSGLVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAIIH 120
QY 136 EVFENSKIYIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIYVRDL 180
| | | | |
DB 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNKPYVGEVDSWSL 240
```

Db 121 EVFENS SKIYIVMEYASRGDLVDYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 180
QY 181 KLENILLDANGNKIKIDFGLSNLYHKGKFLQTCGSPLYASPEIYVNGKPYVGEVDSWSL 240
Db 181 KLENILLDANGNKIKIDFGLSNLYHKGKFLQTCGSPLYASPEIYVNGKPYVGEVDSWSL 240
QY 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLLMWNPTRRATLED 300
Db 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLLMWNPTRRATLED 300
QY 301 VASHMWNVNGYTTGVGEQEARREGHPSGDPGRASADWLRSSRPLLENAGAVCSFFKQ 360
Db 301 VASHMWNVNGYTTGVGEQEARREGHPSGDPGRASADWLRSSRPLLENAGAVCSFFKQ 360
QY 361 HYPGGGSTVPGLERQHSLLKSKRENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
Db 361 HYPGGGSTVPGLERQHSLLKSKRENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
QY 421 TSSEGEVQEDPOELRPVDPDTPGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
Db 421 TSSEGEVQEDPOELRPVDPDTPGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
QY 481 ASDVFSVGPVVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
Db 481 ASDVFSVGPVVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
QY 541 ARSRPSGAVSEDSIISSESFQDLPERLPETPLAGCVSVNLRGLQEPSPSGC---LK 536
Db 541 ARSRPSGAVSEDSIISSESFQDLPERLPETPLAGCVSVNLRGLQEPSPSGC---LK 536
QY 597 RWOESLQDSFSLTDCQEVTAAYRQALGICSKLS 631
Db 597 RWOESLQDSFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 4
US-09-963-159-2
Sequence 2, Application US/09963159
Patent No. US2002007312A1
GENERAL INFORMATION:
APPLICANT: Curtis, Rory A.J.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
FILE REFERENCE: 10147-5011
CURRENT APPLICATION NUMBER: US/09/963,159
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/234,922
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 628
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-159-2

Query Match 85.8%; Score 2824.5; DB 9; Length 628;
Best Local Similarity 85.7%; Pred. No. 4.2e-176;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLOPSPQASALASASARPLADGLIKSPKMKQAVKHHHKNLRHYEFL 60
Db 1 MESLIVARNSGPTPS---AAELARPLABGLIKSPKMKQAVKHHHKNLRHYEFL 56
QY 61 ETLGKGYGVKKARSSGRLVAIKIRKQIKDEODLLIRREIEMSLNPHITAIH 120
Db 57 ETLGKGYGVKKARSSGRLVAIKIRKQIKDEODLLIRREIEMSLNPHITAIH 116
QY 121 EVFENS SKIYIVMEYASRGDLVDYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 180
Db 117 EVFENS SKIYIVMEYASRGDLVDYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 176

QY 181 KLENILLDANGNKIKIDFGLSNLYHKGKFLQTCGSPLYASPEIYVNGKPYVGEVDSWSL 240
Db 177 KLENILLDANGNKIKIDFGLSNLYHKGKFLQTCGSPLYASPEIYVNGKPYVGEVDSWSL 236
QY 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLLMWNPTRRATLED 300
Db 237 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLLMWNPTRRATLED 296
QY 301 VASHMWNVNGYTTGVGEQEARREGHPSGDPGRASADWLRSSRPLLENAGAVCSFFKQ 360
Db 297 VASHMWNVNGYTTGVGEQEARREGHPSGDPGRASADWLRSSRPLLENAGAVCSFFKQ 356
QY 361 HYPGGGSTVPGLERQHSLLKSKRENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
Db 357 HYPGGGSTVPGLERQHSLLKSKRENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 416
QY 421 TSSEGEVQEDPOELRPVDPDTPGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
Db 417 TSSEGEVQEDPOELRPVDPDTPGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 473
QY 481 ASDVFSVGPVVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 540
Db 474 ASDVFSVGPVVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASSHPA 533
QY 541 ARSRPSGAVSEDSIISSESFQDLPERLPETPLAGCVSVNLRGLQEPSPSGC---LK 536
Db 534 ARSRPSGAVSEDSIISSESFQDLPERLPETPLAGCVSVNLRGLQEPSPSGC---LK 533
QY 597 RWOESLQDSFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWOESLQDSFSLTDCQEVTAAYRQALGICSKLS 628

RESULT 5
US-10-423-543-44
Sequence 44, Application US/10423543
Publication No. US20040058355A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Libermann, Rosana K.
APPLICANT: Hunter, John J.
APPLICANT: Meyers, Rachel E.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Olandt, Peter J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Galvin, Katherine M.
APPLICANT: Chun, Miyoung
APPLICANT: Williamson, Mark J.
APPLICANT: Siles-Santiago, Immaculada
TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
FILE REFERENCE: MP103-0230NMIM
CURRENT APPLICATION NUMBER: US/10/423,543
CURRENT FILING DATE: 2003-04-25
PRIOR APPLICATION NUMBER: US 10/278,036
PRIOR FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: US 09/711,216
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/205,447
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 10/012,055
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/248,325
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 10/003,690
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/248,893
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/797,039

```

; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 10/217,168
; PRIOR FILING DATE: 2002-08-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 628
; TYPE: PR1
; ORGANISM: Homo Sapiens
US-10-423-543-44

```

Query Match	85.8%	Score 2824.5	DB 15	Length 628
Best Local Similarity	85.7%	Pred. No. 4,2e-176		
Matches	544	Conservative	26	Mismatches 54; Indels 11; Gaps 3

Qy	1	MESVALLQRPQAPBASALASBSAPPLADGLIKSPKPLMKQAVRRHHKHLRRHYEFL	60
Db	1	MESLVFARRSGPTB---AAELARPLLEGLIKSPKPLMKQAVRRHHKHLRRHYEFL	56
Qy	61	ETLKGATGKVKKARSSGRLVAIYSIRAKDKIKDEODLLHRRLEIYSSLNHPHIIYH	120
Db	57	ETLKGATGKVKKARSSGRLVAIYSIRAKDKIKDEODLLHRRLEIYSSLNHPHIIYH	116
Qy	121	EVFNSSKIVIMESYASRGDLYDYISERPRLSERDARHFQIOYSAHYCHONGIVARDL	180
Db	117	EVFNSSKIVIMESYASRGDLYDYISERQOLSERARHFQIOYSAHYCHONGIVARDL	176
Qy	181	KLKNIILDANANIKIADRFGLSNLYHKGFELTFQCSPLYASPEIYNGKPYGPEVDWSL	240
Db	177	KLKNIILDANANIKIADRFGLSNLYHQGFELTFQCSPLYASPEIYNGKPYGPEVDWSL	236
Qy	241	GVLLYIIVHGTMPPDQDHTLVYKQISNGAYRPEPPDAGCLIKMLIMVNPTRATYED	300
Db	237	GVLLYIIVHGTMPPDGHDKITLVYQISNGAYRPEPPDAGCLIKMLIMVNPTRATYED	296
Qy	301	VASHTWVYVWGTTTGYGEQALREGHGPGDGRASMDWLRSSRPILLENGAKVCSFPKQ	360
Db	297	VASHTWVYVWGATYRGEQEAHPEGHPDSBARASMDWLRSSRPILLENGAKVCSFPKQ	356
Qy	361	HVPGGSTVPLELROHSLKSRKENDMQNLQGDPAEDTSRPGKSSILKPLKGIKKSS	420
Db	357	HAPGGSTTPELEROHSLKSRKENDMQSLHSDPADTAAHFGKSNLKLPRKGIKKYVS	416
Qy	421	TSSGEVQEDPOLRPVPDTPGQPVPAVSLPRKGIKKSRORSRGYSPPBSSEGGELD	480
Db	417	ASAGVQVEDPELRLPIPASPGQAP---LPRKGIKKRPQREGSGYSSPBSSEGGELD	473
Qy	481	ASDVFGSGDPPEOKSPQASGILLHRKGIKKNGKSRFALAGTTSTGSLDQLASSHPA	540
Db	474	AGDVFGSGDPEQKQRPQASGILLHRKGIKKNGKRSQFALBIAAATTGSLDELAPRPL	533
Qy	541	ARPSRPGAVSBDISLSESFDOLDPERLPEPLRGVSYDNLRGLGEQPSGQ---LK	596
Db	534	ARASRPGAVSBDISLSESFDOLDPERLPEPLRGCVSYDNLRGLGEQPSGQCLR	593
Qy	597	RMQBSLGDSCFSITDQCEVTAARQALGICSKUS	631
Db	594	KWRQDPLGDSCTSLDQCEVATYARQALRVCSKLT	628

```

? CURRENT APPLICATION NUMBER: US/10/618,941
? CURRENT FILING DATE: 2003-07-15
? PRIOR APPLICATION NUMBER: 60/395,632
? PRIOR FILING DATE: 2002-07-15
? NUMBER OF SEQ ID NOS: 143
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 77
? LENGTH: 672
? TYPE: prt
? ORGANISM: Homo sapiens
US-10-618-941-77

```

[illegible]

RESULT 6
US-10-618-941-77
Sequence 77, Application US/10618941
Publication No. US20040197792A1
GENERAL INFORMATION:
APPLICANT: WHITE, DAVID
APPLICANT: MANNING, GERRARD
APPLICANT: CAENEPEL, SEAN
TITLE OF INVENTION: NOVEL KINASES
FILE REFERENCE: 034536-0321

RESULT 7
US-10-311-034-12
Sequence 12, Application US/10311103
Publication No. US2004002342A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: YUE, Henry
APPLICANT: IAL, Preeti
APPLICANT: BANMAN, Olga
APPLICANT: BOROMSKI, Mark L.
APPLICANT: LU-YOUNG, Janice
APPLICANT: FU, Yan

```
APPLICANT: GANDHI, Aameena R.
APPLICANT: TRIBOUTLEY, Catherine M.
APPLICANT: CHAMLA, Narinder K.
APPLICANT: YAO, Monique G.
APPLICANT: LU, Dyring Aina M.
APPLICANT: GREENWALD, Sara R.
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: KEARNEY, Liam
APPLICANT: BURFORD, Neil
APPLICANT: NGUYEN, Damien B.
APPLICANT: TRAN, Y. Tom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: HE, Ann
APPLICANT: THORNTON, Michael
APPLICANT: HAPALIA, April
APPLICANT: ARVIZU, Chandra S.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: LO, Terence P.
APPLICANT: KHAH, Farrah A.
APPLICANT: RECIPON, Shirley A.
APPLICANT: AZIMZAI, Yalda
APPLICANT: POLICKY, Jennifer L.
APPLICANT: DING, Li
APPLICANT: GREYHER, Megan
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: BATEA, Sajeev
APPLICANT: ISON, Craig H.
APPLICANT: TITLE OF INVENTION: HUMAN KINASES
APPLICANT: FILE REFERENCE: PI-0125 PCT
APPLICANT: CURRENT APPLICATION NUMBER: US/10/311,034
APPLICANT: CURRENT FILING DATE: 2002-12-10
APPLICANT: PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
APPLICANT: 60/228,056
APPLICANT: PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
APPLICANT: 25
APPLICANT: NUMBER OF SEQ ID NOS: 52
APPLICANT: SOFTWARE: PERL Program
APPLICANT: SEQ ID NO 12
APPLICANT: LENGTH: 594
APPLICANT: TYPE: PRT
APPLICANT: ORGANISM: Homo sapiens
APPLICANT: FEATURE:
APPLICANT: NAME/KEY: misc feature
APPLICANT: OTHER INFORMATION: Inocyte ID No. US20040023242A1 4841542CD1
US-10-311-034-12

Query Match      82.6%; Score 2720.5; DB 15; Length 594;
Best Local Similarity 87.3%; Pred. No. 2,4e-169;
Matches 521; Conservative 22; Mismatches 47; Indels 7; Gaps 2;
```

```
DB 301 WLRSSRPILLENGAKVCSFFKQHPGGSSTTGLERQHSILKKSREKNDMAQSHSTAD 360
DB 399 TSSRPCKSLKCPKGLIKKSSSTSGEVOEDPOELRPVDTQGPVAVSLPRKILK 458
DB 361 TARRCKSLKCPKGLIKKSSSTSGEVOEDPELSPFASQAP---LTPKILK 417
DB 459 SROESGYSSPESSESGELDASDVFSGDPVEOKSPQASGLLRKILKINGKFSRT 518
DB 418 PROESGYSSPESSESGELDAGDVFSGDPVEOKSPQASGLLRKILKINGKFSRT 477
DB 519 ALEGTTSTFGSLDQASHPPARPERPGAVSEDSILSSESPDOLDLPERLPETLRGC 578
DB 478 ALELAPTTFGSLDQASHPPARPERPGAVSEDSILSSESPDOLDLPERLPETLRGC 537
DB 579 VSDVNLRLGLEOPPSG---LKRMOESLGDSCFSLTDCOEYTAAYRQALGICSKLS 631
DB 538 VSDVNLRLGLEOPPSG---LKRMOESLGDSCFSLTDCOEYTAAYRQALGICSKLS 594

RESULT 8
US-10-322-281-26
Sequence 26, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc S. Malandro
APPLICANT: TITLE OF INVENTION: Novel Compositions and Methods in Cancer
APPLICANT: FILE REFERENCE: 529452001000
APPLICANT: CURRENT APPLICATION NUMBER: US/10/322,281
APPLICANT: CURRENT FILING DATE: 2002-12-17
APPLICANT: NUMBER OF SEQ ID NOS: 866
APPLICANT: SOFTWARE: FastSeq for Windows Version 4.0
APPLICANT: SEQ ID NO 26
APPLICANT: LENGTH: 616
APPLICANT: TYPE: PRT
APPLICANT: ORGANISM: Homo sapiens
US-10-322-281-26

Query Match      82.4%; Score 2712.5; DB 16; Length 616;
Best Local Similarity 84.3%; Pred. No. 8.5e-169;
Matches 535; Conservative 25; Mismatches 52; Indels 23; Gaps 9;
```

Db 409 ASAGVQEDPELSPID--GQAAP---LLPKKGIILKKPRQRESGYSSPESESGELLD 463
Qy 481 ASAVFVSQDPVEKQASQASGLLHKRGIILKNGKFSRTALEGTTPTFGSLDOLASHPA 540
Db 464 AGGVFVSQDPVEKQASQASGLLHKRGIILKNGKFSRTALEGTTPTFGSLDOLASHPA 523
Qy 541 ARPSRSGAVSESIISSESFQDLPERLPETPLRCGVVDNLRLGEGPPSG---LK 596
Db 524 ARPSRSGAVSESIISSESFQDLPERLPETPLRCGVVDNLRLGEGPPSGCLR 583
Qy 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 584 RWRQDPLGDSGCSFSLTDCQEVTA--YQALRVCSKLT 616

RESULT 9
US-10-363-829-404
; Sequence 404, Application US/10363829
; Publication No. US20040142331A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;
; APPLICANT: Dahl, Christopher R.; Momiyama, Monika G.;
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;
; APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyna H.;
; APPLICANT: David, Marie H.; Panzer, Scott R.;
; APPLICANT: Flores, Vincent Z.; Dafio, Abel;
; APPLICANT: Marwaha, Rakesh; Chen, Alice U.;
; APPLICANT: Chang, Simon C.; Au, Alan P.;
; APPLICANT: Imman, Rebekah R.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1183 USN
; CURRENT APPLICATION NUMBER: US/10/363,829
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: PCT/US01/27628
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/229,751
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,749
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,750
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,747
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,748
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,583
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,517
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,610
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,597
; PRIOR FILING DATE: 2000-09-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PERL Program
; SEQ ID NO 404
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: LG:982800.1.orfi:2000SEP08
US-10-363-829-404

Query Match 54.1%; Score 1781.5; DB 16; Length 406;
Best Local Similarity 83.6%; Pred. No. 3.3e-108;

Matches 342; Conservative 16; Mismatches 44; Indels 7; Gaps 2;
Qy 227 GRRYVPEVDWSLGLVLYLVHGTMPFDGQDKTLVKQISNGAYREPPKPSDACGLIRW 286
Db 1 GRRYTPPEVDWSLGLVLYLVHGTMPFDGQDKTLVKQISNGAYREPPKPSDACGLIRW 60
Qy 287 LLMVNTTRRAATLEDAVSHMMVMNGYTTGGEQALREGGHPSDFRASMADLRRSSRP 346
Db 61 LLMVNTTRRAATLEDAVSHMMVMNGYATRVGEQALREGGHPSDSARASADLRRSSRP 120
Qy 347 LLENAGKVSFPFROHVPGGGASTVPLGEROHSILKSKRENDMAQNLGDPADERTSSRPGKS 406
Db 121 LLENAGKVSFPFROHVPGGGASTVPLGEROHSILKSKRENDMAQNLGDPADERTSSRPGKS 180
Qy 407 SLKLPKGIILKKKSTSSGVEQEDPELRPVPTPGQPVAVSLPPKGIILKKSRQESGY 466
Db 181 NLKLPKGIILKKKVSASAGVQEDPELSPIDPASPQGAAP---LLPKKGIILKKRQESGY 227
Qy 467 YSSPESSEGGELLDADVFSVSDPVQKSPQASGLLHKRGIILKNGKFSRTALEGTTPT 526
Db 238 YSSPESSEGGELLDADVFSVSDPVQKSPQASGLLHKRGIILKNGKFSRTALEGTTPT 297
Qy 527 TFGSLDOLASHPAARPSRSGAVSESIISSESFQDLPERLPETPLRCGVVDNLRLG 586
Db 298 TFGSLDOLASHPAARPSRSGAVSESIISSESFQDLPERLPETPLRCGVVDNLRLG 357
Qy 587 LRQPPSEG---LKRWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 358 LRQPPSEGSGCLRRWRQDPLGDSGCSFSLTDCQEVTAAYRQALRVCSKLT 406

RESULT 10
US-10-363-829-504
; Sequence 504, Application US/10363829
; Publication No. US20040142331A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;
; APPLICANT: Dahl, Christopher R.; Momiyama, Monika G.;
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;
; APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyna H.;
; APPLICANT: David, Marie H.; Panzer, Scott R.;
; APPLICANT: Flores, Vincent Z.; Dafio, Abel;
; APPLICANT: Marwaha, Rakesh; Chen, Alice U.;
; APPLICANT: Chang, Simon C.; Au, Alan P.;
; APPLICANT: Imman, Rebekah R.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1183 USN
; CURRENT APPLICATION NUMBER: US/10/363,829
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: PCT/US01/27628
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/229,751
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,749
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,750
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,747
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,748
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,583
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,517
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,610
; PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230,597
PRIOR FILING DATE: 2000-09-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 506
SOFTWARE: PERL Program
SEQ ID NO 504
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incycle ID No: LI:1144409.1.orfl:2000SEP08
US-10-363-829-504

Query Match 54.1%; Score 1781.5; DB 16; Length 406;
Best Local Similarity 83.6%; Pred. No. 3,3e-108;
Matches 342; Conservative 16; Mismatches 44; Indels 7; Gaps 2;

QY 227 GKPYVPEVDWSGLVLYLVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDAGLIRW 286
DB 1 GKPYTPEVDWSGLVLYLVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDAGLIRW 60
QY 287 LLMVNTTRATLEDVASHMMVWNGYTTGVGEQALRGGHPSGDFGASADWLRRSSRP 346
DB 61 LLMVNTTRATLEDVASHMMVWNGYTTGVGEQALRGGHPSGDFGASADWLRRSSRP 120
QY 347 LLENGAKVCSFFKQHPVGGGSTVGLERHSLKSRKENDMAONLQGDPAEDTSSRPGKS 406
DB 121 LLENGAKVCSFFKQHPVGGGSTVGLERHSLKSRKENDMAONLQGDPAEDTSSRPGKS 180
QY 407 SLKPLKGLIKKKSSTSGEVEDPOELRPVDTPGQVPVAVSLPRKGLIKKSRORESGY 466
DB 181 NLKPLKGLIKKKSSTSGEVEDPOELRPVDTPGQVPVAVSLPRKGLIKKSRORESGY 237
QY 467 YSSPESSEGEELDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTALRSTPS 526
DB 238 YSSPESSEGEELDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTALRSTPS 297
QY 527 TFGSLDQLASHPARPSPRGAVSEDSILSSSPDLDLPERLPETPLNGCVSYNDLRG 586
DB 298 TFGSLDQLASHPARPSPRGAVSEDSILSSSPDLDLPERLPETPLNGCVSYNDLRG 357
QY 587 LEOPPESEG---LKRWQBSLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 358 LEOPPESEGSCLRWRQDPLGDSCSFSLTDCQEVTAAYRQALGICSKLT 406

RESULT 11
US-09-780-949-2
Sequence 2, Application US/09780949

Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Weich, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
TITLE OF INVENTION: Kinase
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 661
TYPE: PRT
ORGANISM: H. sapiens
US-09-780-949-2

Query Match 51.6%; Score 1699.5; DB 9; Length 661;
Best Local Similarity 55.5%; Pred. No. 1.4e-102;
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

QY 14 PSASALASSAPRLADGLIKSPRLMKQAVKHHKHNIRHYELETIGKTYGKVK 73
DB 13 PDGLGAPSPREAVAGATPALEP-RKPHGVKHHKHNIRHYELETIGKTYGKVK 71
QY 74 ARSS-SGRVATKSTIKDKIKODLHTRREIEMSSLNHPHIIHVEFESSKVIY 132
DB 72 ATERFSGRVATKSTIKDKIKODLHTRREIEMSSLNHPHIIHVEFESSKVIY 131
QY 133 MEYASRGDLYDYSERPRLSERDARHPFQIVSALYCHONGIYVHDLKLENTILDANGN 192
DB 132 MEYASRGDLYDYSERPRLSERDARHPFQIVSALYCHONGIYVHDLKLENTILDANGN 191
QY 193 IKIADFGSLNLYKRGFLQTFCCSPLYASBEIYNGKRYVPEVDWSGLVLYLVHGTM 252
DB 192 IKIADFGSLNLYKRGFLQTFCCSPLYASBEIYNGKRYVPEVDWSGLVLYLVHGTM 251
QY 253 PFDGQDHKTLVKQISNGAYRPPKPSDAGLIRWLLMVNTTRATLEDVASHMMVWNGY 312
DB 252 PFDGQDHKTLVKQISNGAYRPPKPSDAGLIRWLLMVNTTRATLEDVASHMMVWNGY 311
QY 313 TGVGEQALRGGHPSGDFGASADWLRRSSRPLLENGAKVCSFFKQHPVGGGSTVGL 372
DB 312 TGVGEQALRGGHPSGDFGASADWLRRSSRPLLENGAKVCSFFKQHPVGGGSTVGL 362
QY 373 ERQSLKSRKENDMAONLQGDPAEDTSSRPGKSILKPKGLIKKS-----STSGEVQ 427
DB 363 ERQSLKSRKENDMAONLQGDPAEDTSSRPGKSILKPKGLIKKS-----STSGEVQ 418
QY 428 -----EDPDEL-----RPVDTPGQVPVAVS-LPRKGLIKKSRORESGY 466
DB 419 GVVPALPSTFKMEQDLCRTGVLLPSSPEAEVPGKLSPKQASATMPKGLIKKQOORESGY 478
QY 467 YSSPESSEGEELDADVDVPSGDPVBOKSP-----QASGLLHRKGLIKLNGKESRTAL 520
DB 479 YSSPESSEGEELDADVDVPSGDPVBOKSP-----QASGLLHRKGLIKLNGKESRTAL 536
QY 521 EGTTPS---TFGSLDQ-----LASHPARPSPRGAVSEDSILSSSPDLDLPER 569
DB 537 DPAVPEPMTLESLEPGVAPAGLSRY-----SRPSSVTSIDSVLSSGFDLDDIQEN 591
QY 570 LP-ETPLRGCVSYNDL-----LRLGLOPPESEGSLKRWQBSLGDSCFS-LTDCQEVTAAYR 621
DB 592 RPARQRIKSCVSAENFLQIDFGLQNRPPQYIKRYRNLADSSFSLLTDMDDVTGVYK 651
QY 622 QALGICSKLS 631
DB 652 QALGICSKLN 661

RESULT 12

US-09-780-949-6
Sequence 6, Application US/09780949
Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Weich, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
TITLE OF INVENTION: Kinase
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 661
TYPE: PRT
ORGANISM: H. sapiens
US-09-780-949-6

```

1 TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
2 FILE REFERENCE: MPI02-020P1ROMNITM
3 CURRENT APPLICATION NUMBER: US/10/354,358
4 CURRENT FILING DATE: 2003-01-30
5 PRIOR APPLICATION NUMBER: US 60/353,600
6 PRIOR FILING DATE: 2002-01-31
7 PRIOR APPLICATION NUMBER: US 60/364,517
8 PRIOR FILING DATE: 2002-03-15
9 PRIOR APPLICATION NUMBER: US 60/371,075
10 PRIOR FILING DATE: 2002-04-09
11 PRIOR APPLICATION NUMBER: US 60/371,507
12 PRIOR FILING DATE: 2002-04-10
13 PRIOR APPLICATION NUMBER: US 60/372,984
14 PRIOR FILING DATE: 2002-04-16
15 PRIOR APPLICATION NUMBER: US 60/374,194
16 PRIOR FILING DATE: 2002-04-19
17 PRIOR APPLICATION NUMBER: US 60/382,995
18 PRIOR FILING DATE: 2002-05-24
19 PRIOR APPLICATION NUMBER: US 60/385,023
20 PRIOR FILING DATE: 2002-05-31
21 PRIOR APPLICATION NUMBER: US 60/388,853
22 PRIOR FILING DATE: 2002-06-11
23 PRIOR APPLICATION NUMBER: US 60/389,395
24 PRIOR FILING DATE: 2002-06-17
25 Remaining Prior Application data removed - See file wrapper or PALM.
26 NUMBER OF SEQ ID NOS: 122
27 SOFTWARE: FastSeq for Windows Version 4.0
28 SEQ ID NO 82
29 LENGTH: 661
30 TYPE: PRT
31 ORGANISM: Homo sapiens
32 US-10-354-358-82
33
34 Query Match 51.6%; Score 1699.5; DB 14; Length 661;
35 Best Local Similarity 55.5%; Pred. No. 14e-102;
36 Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;
37
38 14 PSASALASASAPLADGLIKSPKPMKKOAVRRHHKNNLRHYEFLETGKGTGKVKK 73
39 13 PDLGAGAGSREAVAGATALLP-RKPGVGRHHKNNLKRHYELQFTLGKGTGKVKR 71
40 74 ARBS-SGRVAIKSIRKDKIKDEODLHIREIETMSSLNHPHITAIHEVENSKIYIV 132
41 72 ATERPFGGVVAIKSIRKDKIKDEODMVIHREIEIMSSLNHPHIIISIEVEFNKDKVII 131
42 133 MEVSRGDLVYISRRPRLSERDARHFRQIVSAHYCHQNGIYHNDLKLIENTLDANGN 132
43 132 MEVASKGGLYIISERRRLESEKETHFRQIVSAHYCHKNGVHNDLKLIENTLDNCN 131
44 193 IKIADFGLSNLYHKKFLQTFGCSPLVYASPEIYNCKPVYGVPEVDSWLSGLVLYILVHGTM 252
45 192 IKIDFGLSNLVQXDKFQTFGCSPLVYASPEIYNRPYGPPEVDSALGLVLYLYVYGTM 251
46 253 PFDGQDHTLVKQISNGAYREBPKPSDACGLIRMLMNVPTRRATLEDVASHMWNVWGYT 312
47 252 PFDGFDHKNLIRQISSLGSEYREPTQSDARGLIRMLMNVPTRRATLEDIANHWMVWVWGYX 311
48 313 TGVBGEGLLRGCGHSGDPCGASMDWLRSSRPLENGAKVCSFPKQHYVPGGSGTVPGU 372
49 312 SSVDCDCLADHSESLP-LARIIDMHRSIGLADTEANKGLAKL--PTTSEVM-L 362
50 373 ERQSLKRSKRENDMAQNLQGDPAEDTSRPPKSSKLKPKGLIKKKSS-----STSGEAV 427
51 363 ERQSLKRSKRENDMAQNLQGDPAEDTSRPPKSSKLKPKGLIKKKSS-----STSGEAV 418
52 428 -----EDPQEL-----RPVPTPGQVPVAVS-LPRLKGLIKKRSRQRESG 466
53 419 GVPAPALPSTPKMEQDLCRTGVLLPSSPEAEVPGTLSPQASATWPKGLIKKTKQRESG 478
54 467 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 520
55 479 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
56 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
57 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
58 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
59 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
60 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
61 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
62 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
63 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
64 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
65 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
66 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
67 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
68 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
69 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
70 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
71 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
72 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
73 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
74 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
75 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
76 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
77 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
78 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
79 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
80 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
81 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
82 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
83 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
84 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
85 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
86 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
87 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
88 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
89 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
90 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
91 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
92 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
93 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
94 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
95 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
96 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
97 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
98 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
99 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516
100 YSSPEPSGSLDASDVYFVSGDPVPEQKSP-----QASGLLHRKGLIKLKGKESRTAL 516

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 22:46:12 ; Search time 1080.14 Seconds
(without alignments)
10726.042 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atgagatcggtgagcttact.....gaatctgtcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	16 US-10-355-975-4	Sequence 4, Appl1
2	1893	100.0	3073	18 US-10-322-281-22	Sequence 22, Appl1
3	1888.2	99.7	2027	18 US-10-343-514-87	Sequence 87, Appl1
4	1683.4	88.9	2026	18 US-10-343-514-27	Sequence 27, Appl1
5	1683.4	88.9	2929	18 US-10-343-514-1	Sequence 1, Appl1
6	1360.2	71.9	1884	9 US-09-963-159-3	Sequence 3, Appl1
7	1360.2	71.9	1884	17 US-10-423-543-45	Sequence 45, Appl1
8	1360.2	71.9	3353	9 US-09-963-159-1	Sequence 1, Appl1
9	1360.2	71.9	3353	17 US-10-423-543-43	Sequence 43, Appl1
10	1357.6	71.7	3360	17 US-10-311-034-38	Sequence 38, Appl1
11	1357	71.7	3404	18 US-10-322-281-25	Sequence 25, Appl1

12	1357	71.7	3463	18 US-10-618-941-11	Sequence 11, Appl1
13	1355.4	71.6	2501	17 US-10-302-172-215	Sequence 215, Appl1
14	1355.4	71.6	3443	18 US-10-370-715B-639	Sequence 639, Appl1
15	1071.4	56.6	1186	18 US-10-343-514-13	Sequence 13, Appl1
16	1059.4	56.0	37278	18 US-10-322-281-21	Sequence 21, Appl1
17	943.4	49.8	1166	18 US-10-343-514-70	Sequence 70, Appl1
18	840.2	44.4	2616	18 US-10-363-829-151	Sequence 151, Appl1
19	840.2	44.4	2619	18 US-10-363-829-250	Sequence 250, Appl1
20	723	38.2	39699	18 US-10-322-821-24	Sequence 24, Appl1
21	648	34.2	707	18 US-10-343-514-91	Sequence 91, Appl1
22	648	34.2	712	18 US-10-343-514-31	Sequence 31, Appl1
23	644.2	34.0	661	18 US-10-343-514-93	Sequence 93, Appl1
24	613.6	32.4	741	18 US-10-343-514-32	Sequence 32, Appl1
25	592	31.3	734	18 US-10-343-514-92	Sequence 92, Appl1
26	583.2	30.8	588	18 US-10-343-514-97	Sequence 97, Appl1
27	557.4	29.4	660	18 US-10-343-514-33	Sequence 33, Appl1
28	557.4	28.2	578	18 US-10-343-514-95	Sequence 95, Appl1
29	515	27.2	548	18 US-10-343-514-96	Sequence 96, Appl1
30	510.6	27.0	579	18 US-10-343-514-35	Sequence 35, Appl1
31	506.4	26.8	1986	9 US-09-780-949-3	Sequence 3, Appl1
32	506.4	26.8	2884	9 US-09-780-949-5	Sequence 5, Appl1
33	506.4	26.8	6828	9 US-09-780-949-1	Sequence 1, Appl1
34	506.4	26.8	6828	16 US-10-354-358-81	Sequence 81, Appl1
35	506.4	26.8	6828	17 US-10-172-118-1541	Sequence 1541, Appl1
36	506.4	26.8	6828	17 US-10-342-887-1541	Sequence 1541, Appl1
37	506.4	26.8	6828	18 US-10-723-860-2552	Sequence 2552, Appl1
38	506.4	26.8	6854	18 US-10-723-860-6703	Sequence 6703, Appl1
39	493.6	26.1	585	18 US-10-343-514-37	Sequence 37, Appl1
40	493.6	25.6	548	18 US-10-343-514-36	Sequence 36, Appl1
41	448.2	23.7	521	18 US-10-343-514-94	Sequence 94, Appl1
42	417.8	22.1	521	18 US-10-343-514-34	Sequence 34, Appl1
43	383.4	20.3	439	18 US-10-343-514-10	Sequence 10, Appl1
44	381.4	20.1	436	18 US-10-343-514-14	Sequence 14, Appl1
45	365.6	19.3	421	18 US-10-343-514-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-10-355-975-4
; Sequence 4, Application US/10355975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/10/355,975
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579,664B
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-355-975-4

Query Match 100.0%; Score 1893; DB 16; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 123 ATGAGTCGGTGGCTTACTTCACAGCCGAGCAGGCTCCCTGCGCTCGCCCTGAGC 60
1 ATGAGTCGGTGGCTTACTTCACAGCCGAGCAGGCTCCCTGCGCTCGCCCTGAGC 60
61 TCGAGAGCCGCCGCGCTGGCGAGCGGCTCATCAAGTCGCTTAACCTTGATGAAG 120
TCGAGAGCCGCCGCGCTGGCGAGCGGCTCATCAAGTCGCTTAACCTTGATGAAG 120

Db 183 TCGAGAGCCGCCGCCGCTGCGGAGCGGCTCATCAAGTCGCTAAACCTGTATAG 242
QY 121 AAGCAGCCGCTGAAGCGGCACATCAACAACAACCTCGCGCACCCGTACGATTCCTG 180
Db 243 AACCAAGCCGGTGAAGCCGACCATCAACAACAACCTCGCGCACCCGTACGATTCCTG 302
QY 181 GAGACGCTGGGCAAGGCACTACGCGGAGGTGAAGAAGGACGAGAGAGCTGGGGCGT 240
Db 303 GAGACGCTGGGCAAGGCACTACGCGGAGGTGAAGAAGGACGAGAGAGCTGGGGCGT 362
QY 241 CTGGTGGCCATCAAGTCATCAGGAAACAAAATCAAGATGACAGATCTGCTGAC 300
Db 363 CTGGTGGCCATCAAGTCATCAGGAAACAAAATCAAGATGACAGATCTGCTGAC 422
QY 301 ATACGAGGAGATGAGATCATCTTCACTCAACCAACCCCACTCATATTCATCAT 360
Db 423 ATACGAGGAGATGAGATCATCTTCACTCAACCAACCCCACTCATATTCATCAT 482
QY 361 GAAGTGTGAGATACAGCAAGATGATGATGTCATGAGATGTCAGCCGAGGAGAT 420
Db 483 GAAGTGTGAGATACAGCAAGATGATGATGTCATGAGATGTCAGCCGAGGAGAT 542
QY 421 CTGTATGATTAATCAATCAATGAGCGGCAACGCGTGAAGAGCGGAGCATTTCTTC 480
Db 543 CTGTATGATTAATCAATCAATGAGCGGCAACGCGTGAAGAGCGGAGCATTTCTTC 602
QY 481 CGACAGATCGTGTGCTGCTGCACTACTGCAACGAGATGTCATCCGAGATCTC 540
Db 603 CGACAGATCGTGTGCTGCTGCACTACTGCAACGAGATGTCATCCGAGATCTC 662
QY 541 AAGCTGGAACCATCTTCTGATGCAATGGAACCATCAAGATGTCATTTGGGCTC 600
Db 663 AAGCTGGAACCATCTTCTGATGCAATGGAACCATCAAGATGTCATTTGGGCTC 722
QY 601 TCCAACTGTACCAACAAGGCAAGTTCCTCAGACGTTCTGTGGAGACCCCTCTCAAGC 660
Db 723 TCCAACTGTACCAACAAGGCAAGTTCCTCAGACGTTCTGTGGAGACCCCTCTCAAGC 782
QY 661 TCCGCTGAGATGATCAACGCGGAAGCCCTATGTGGGCCCAAGATGGAACGCTGCTG 720
Db 783 TCCGCTGAGATGATCAACGCGGAAGCCCTATGTGGGCCCAAGATGGAACGCTGCTG 842
QY 721 GGGGTTCTCCTGATCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 843 GGGGTTCTCCTGATCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 902
QY 781 ACACTGATGAAGAAATCAATTAACGCGGCTTACCGTAGCGCGCAAGCGTCCGATGCC 840
Db 903 ACACTGATGAAGAAATCAATTAACGCGGCTTACCGTAGCGCGCAAGCGTCCGATGCC 962
QY 841 TGTGGCTGTATCCGATGCTGTTATGATGATGATGATGATGATGATGATGATGAT 900
Db 963 TGTGGCTGTATCCGATGCTGTTATGATGATGATGATGATGATGATGATGATGAT 1022
QY 901 GTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1023 GTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 961 CTGCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1083 CTGCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
QY 1021 CGTGGCTCTCGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1143 CGTGGCTCTCGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1202
QY 1081 CAGTGTCCGGGAGTGAAGCACTGTACCTGGGTGAGAGCGGCAACTTCTTTAAGAG 1140
Db 1203 CAGTGTCCGGGAGTGAAGCACTGTACCTGGGTGAGAGCGGCAACTTCTTTAAGAG 1262
QY 1141 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1263 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322

QY 1201 TCTGCGCTTGGCAAGACACCTTTAAGCTTCCGAAAGGCAATTCACAAAGAAAGTCTCT 1260
Db 1323 TCTGCGCTTGGCAAGACACCTTTAAGCTTCCGAAAGGCAATTCACAAAGAAAGTCTCT 1382
QY 1261 ACCCTGTCAGGAGGATGACAGAGGACCTCAGGAACTCAGACCGGCTGATCTCA 1320
Db 1383 ACCCTGTCAGGAGGATGACAGAGGACCTCAGGAACTCAGACCGGCTGATCTCA 1442
QY 1321 GGGGAGCTGTCTCCCTGTATCTCTGCTCCAGGAAAGGCAATCTTAAAGATCTCA 1380
Db 1443 GGGGAGCTGTCTCCCTGTATCTCTGCTCCAGGAAAGGCAATCTTAAAGATCTCA 1502
QY 1381 CAGCGTAAATCTGTTACTACTCTCTCCAGAGCCAGGAGATCTGGGAACTTTAGAC 1440
Db 1503 CAGCGTAAATCTGTTACTACTCTCTCCAGAGCCAGGAGATCTGGGAACTTTAGAC 1562
QY 1441 GCCAGTAAATGTTTGAATGAGGAGCCCGGTGAGAGGATCTCCAGAGGCTTCAGG 1500
Db 1563 GCCAGTAAATGTTTGAATGAGGAGCCCGGTGAGAGGATCTCCAGAGGCTTCAGG 1622
QY 1501 CTCTCTCTCCAGGCAAGGCAATTTCAAACTCAATGCAAGTTCCTCCGACAGCTTA 1560
Db 1623 CTCTCTCTCCAGGCAAGGCAATTTCAAACTCAATGCAAGTTCCTCCGACAGCTTA 1682
QY 1561 GAAGGCACTACCTCTAGACCTTTGGCTCTGAGACCACTGACCTCTCCATCTGCA 1620
Db 1683 GAAGGCACTACCTCTAGACCTTTGGCTCTGAGACCACTGACCTCTCCATCTGCA 1742
QY 1621 GCCCGGCGCAGCGCCCTCAGGAGGCTGTAAGGAGGACGACATCTGCTCCGATGC 1680
Db 1743 GCCCGGCGCAGCGCCCTCAGGAGGCTGTAAGGAGGACGACATCTGCTCCGATGC 1802
QY 1681 TTTGACCAATTTGACTTGTGCTGTAAGCTTCCGAAACCCCACTGAGGAGCTGTGCT 1740
Db 1803 TTTGACCAATTTGACTTGTGCTGTAAGCTTCCGAAACCCCACTGAGGAGCTGTGCT 1862
QY 1741 GTGAGCAACTGAGGAGGCTTGAAGAGCTCTCTCAGAGGTTTGAAGCATGATGAG 1800
Db 1863 GTGAGCAACTGAGGAGGCTTGAAGAGCTCTCTCAGAGGTTTGAAGCATGATGAG 1922
QY 1801 GAATCCTTGGGAGATGAGCTTTTCTGACAGACCTGCAAGAGGATGATGAGCTTAC 1860
Db 1923 GAATCCTTGGGAGATGAGCTTTTCTGACAGACCTGCAAGAGGATGATGAGCTTAC 1982
QY 1861 AGACAAGCCCTAGAAATCTGCTCAAGCTCAGC 1893
Db 1983 AGACAAGCCCTAGAAATCTGCTCAAGCTCAGC 2015

RESULT 2
US-10-322-281-22
; Sequence 22, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201000
; CURRENT APPLICATION NUMBER: US/10/322,281
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 3073
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-322-281-22

Query Match 100.0%; Score 1893; DB 18; Length 3073;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTGGCTGCTTACTTCAGCGCCCGAGCCAGGCTCTGCGCTCCGCTGACC 60
DB 92 ATGAGTGGCTGCTTACTTCAGCGCCCGAGCCAGGCTCTGCGCTCCGCTGACC 151
QY 61 TCGGAGGCGCCCGCTGCGGAGGCGCTCAAGTGCCTTAACCTCTGATAG 120
DB 152 TCGGAGGCGCCCGCTGCGGAGGCGCTCAAGTGCCTTAACCTCTGATAG 211
QY 121 AAGCAGCGGTGAGCGGCACTCAACACCAACCTGCGGCACTGATTCCTG 180
DB 212 AAGCAGCGGTGAGCGGCACTCAACACCAACCTGCGGCACTGATTCCTG 271
QY 181 GAGAGCTGCGGCAAGGCACTCAAGGAGTGAAGAGGCAAGAGAGCTCGGCGCT 240
DB 272 GAGAGCTGCGGCAAGGCACTCAAGGAGTGAAGAGGCAAGAGAGCTCGGCGCT 331
QY 241 CTGCTGCGCATCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 332 CTGCTGCGCATCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391
QY 301 ATAGCAGGAGAGATGAGATCATCTTCACTCAACCACTCCCATCATCTGCGCAT 360
DB 392 ATAGCAGGAGAGATGAGATCATCTTCACTCAACCACTCCCATCATCTGCGCAT 451
QY 361 GAACTGTTGAGATGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 420
DB 452 GAACTGTTGAGATGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 511
QY 421 CTGATATATTAATCAATGAGTGAAGGCGGCAAGGCTGAGTGAAGGCGGCAAG 480
DB 512 CTGATATATTAATCAATGAGTGAAGGCGGCAAGGCTGAGTGAAGGCGGCAAG 571
QY 481 CGACAGATGAGTGAAGGCGGCAAGGCTGAGTGAAGGCGGCAAGGCTGAGTGA 540
DB 572 CGACAGATGAGTGAAGGCGGCAAGGCTGAGTGAAGGCGGCAAGGCTGAGTGA 631
QY 541 AAGCTGAGAGAGATCTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 632 AAGCTGAGAGAGATCTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 691
QY 601 TCCAGCTGTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 692 TCCAGCTGTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
QY 661 TCCAGCTGTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 752 TCCAGCTGTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
QY 721 GGGCTTCTCTGTAACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 812 GGGCTTCTCTGTAACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871
QY 781 ACACTGTAAGCAATCAAGTGAAGGCGGCTTACCTGAGAGCGGCAAGCGCTG 840
DB 872 ACACTGTAAGCAATCAAGTGAAGGCGGCTTACCTGAGAGCGGCAAGCGCTG 931
QY 841 TGTGCGCTGATCCGCTGCTGTTAATGATGAACCCCACTGCGGCGCACTGAGAG 900
DB 932 TGTGCGCTGATCCGCTGCTGTTAATGATGAACCCCACTGCGGCGCACTGAGAG 991
QY 901 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 992 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1051
QY 961 CTGCTGAG 1020
DB 1052 CTGCTGAG 1111
QY 1021 CGTGCCTCTCGGCGCCCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1112 CGTGCCTCTCGGCGCCCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
QY 1081 CAGCTGCGGAG 1140

DB 1172 CAGCTGCGGAG 1231
QY 1141 TCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1232 TCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
QY 1201 TCTGCGCTGAG 1260
DB 1292 TCTGCGCTGAG 1351
QY 1261 ACTCTGTCAGGAG 1320
DB 1352 ACTCTGTCAGGAG 1411
QY 1321 GGGAGAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1412 GGGAGAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
QY 1381 CAGCTGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1472 CAGCTGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1531
QY 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1532 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
QY 1501 CTCTCTCTCAAGGAG 1560
DB 1592 CTCTCTCTCAAGGAG 1651
QY 1561 GAAAGCAGTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1652 GAAAGCAGTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711
QY 1621 GCCCGGCGCAGCGCGCCCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1712 GCCCGGCGCAGCGCGCCCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
QY 1681 TTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1772 TTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1831
QY 1741 GTGAGCAACTGAG 1800
DB 1832 GTGAGCAACTGAG 1891
QY 1801 GAATCTTGGGGGATAGCTCTTTCTGACAGACTGCAAGAGTGAAGTGAAGCTTAC 1860
DB 1892 GAATCTTGGGGGATAGCTCTTTCTGACAGACTGCAAGAGTGAAGTGAAGCTTAC 1951
QY 1861 AGACAGAGCTGAGAGATGCTCAAGAGTCAAG 1893
DB 1952 AGACAGAGCTGAGAGATGCTCAAGAGTCAAG 1984

RESULT 3
US-10-343-514-87
; Sequence 87, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl P.
; APPLICANT: LEBEVRE, Diana L.
; TITLE OR INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613

; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87
; LENGTH: 2027
; TYPE: DNA
; ORGANISM: MOUSE
US-10-343-514-87

Query Match 99.7%; Score 1888.2; DB 18; Length 2027;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1890; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATGAGATCGGTGGCTTAACTTCAAGCCGCCGAGCCAGGCTCCCTGGGCTCCGCCCTGGCC	60
Db	75	ATGAGATCGGTGGCTTAACTTCAAGCCGCCGAGCCAGGCTCCCTGGGCTCCGCCCTGGCC	134
Qy	61	TCCGAGAGCCGCCGCTGGCGGAGCGGCTCATCAAGTCCCTAAACCTGTATGAG	120
Db	135	TCCGAGAGCCGCCGCTGGCGGAGCGGCTCATCAAGTCCCTAAACCTGTATGAG	194
Qy	121	AAGCAGCGGTGAAGCGGCACTACAAACAACCTGCGGCAACGCTAAGATTCTG	180
Db	195	AAGCAGCGGTGAAGCGGCACTACAAACAACCTGCGGCAACGCTAAGATTCTG	254
Qy	181	GAGACGCTGGGCAAGGCACTTACGGGAGGTGAAGAGGACGAGAGCTCGGGGCT	240
Db	255	GAGACGCTGGGCAAGGCACTTACGGGAGGTGAAGAGGACGAGAGCTCGGGGCT	314
Qy	241	CTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGAGGATCTGCTGAC	300
Db	315	CTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGAGGATCTGCTGAC	374
Qy	301	ATAAGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT	360
Db	375	ATAAGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT	434
Qy	361	GAAATGTTTGAAGATACGAGAAATTTGTATTTGTATGAGATTTGCCAGCGGCGAT	420
Db	435	GAAATGTTTGAAGATACGAGAAATTTGTATTTGTATGAGATTTGCCAGCGGCGAT	494
Qy	421	CTGTATGATTTACATCAGTGAAGCGGCAACGCTGAGTGAAGGAGCGGCAATTTCTTC	480
Db	495	CTGTATGATTTACATCAGTGAAGCGGCAACGCTGAGTGAAGGAGCGGCAATTTCTTC	554
Qy	481	CGACAGATCGTGTGCTGCTGCACTACTGCGCACCAAGACGCGATCTTCAACGAGATCTC	540
Db	555	CGACAGATCGTGTGCTGCTGCACTACTGCGCACCAAGACGCGATCTTCAACGAGATCTC	614
Qy	541	AAGCTGAAAACATCTCTTATGATGCAATGAAAACATCAAGATTGCTGACTTTGGCCTC	600
Db	615	AAGCTGAAAACATCTCTTATGATGCAATGAAAACATCAAGATTGCTGACTTTGGCCTC	674
Qy	601	TCCAACTGTACCAAAAGGCAAGTTCTCCAGAGTTCTGTGGGAGCCCTCTTAAGCC	660
Db	675	TCCAACTGTACCAAAAGGCAAGTTCTCCAGAGTTCTGTGGGAGCCCTCTTAAGCC	734
Qy	661	TGCGCTGAGATGATCAACGAGAAACCTATGTGGGCCCAAGAGTGAACAGTGTCTCTG	720
Db	735	TGCGCTGAGATGATCAACGAGAAACCTATGTGGGCCCAAGAGTGAACAGTGTCTCTG	794
Qy	721	GGGCTTCTCTGTATCATCTGTGTGATGCAACATGCCCCCTTTGAACGAGGAGATCAATAA	780
Db	795	GGGCTTCTCTGTATCATCTGTGTGATGCAACATGCCCCCTTTGAACGAGGAGATCAATAA	854
Qy	781	ACACTGTGAAGCAATCAATGAACGAGGCTTTAACGTTAGCCGCCCAAGCGCTCCGAGGCC	840
Db	855	ACACTGTGAAGCAATCAATGAACGAGGCTTTAACGTTAGCCGCCCAAGCGCTCCGAGGCC	914
Qy	841	TGTGGCTGTATCCGGTGGCTGTTAATGTGAACCCCAACCCGTCGGGCACTGAGAGAT	900

Db	915	TGTGGCTGTATCCGGTGGCTGTTAATGTGAACCCCAACCCGTCGGGCACTGAGAGAT	974
Qy	901	GTAAGCAGTATTTGGTGGGCACTGGGCTTAAACCAACCGGAGTGGGGAACAGGAAGCC	960
Db	975	GTAAGCAGTATTTGGTGGGCACTGGGCTTAAACCAACCGGAGTGGGGAACAGGAAGCC	1034
Qy	961	CTGCGTGAAGGTGGGCACTCTAGTGTGATCTTTGGCGGGGCTTCATGCGGAGTGTGA	1020
Db	1035	CTGCGTGAAGGTGGGCACTCTAGTGTGATCTTTGGCGGGGCTTCATGCGGAGTGTGA	1094
Qy	1021	CGTGGCTCTCGGCGCCCTCTCTGAGATGAGCCCAAGGTGTGACGCTTCTTCAACAG	1080
Db	1095	CGTGGCTCTCGGCGCCCTCTCTGAGATGAGCCCAAGGTGTGACGCTTCTTCAACAG	1154
Qy	1081	CAGTGTCCGAGAGTGTGAACATCTATCTGTGGCTGGAAGCGGCAATCTTCTTAAGAG	1140
Db	1155	CAGTGTCCGAGAGTGTGAACATCTATCTGTGGCTGGAAGCGGCAATCTTCTTAAGAG	1214
Qy	1141	TCCGAAAGGAGATGATCATGAGCTCAAAATCTGCAAGGTGACCCGCGTGAATACCTCT	1200
Db	1215	TCCGAAAGGAGATGATCATGAGCTCAAAATCTGCAAGGTGACCCGCGTGAATACCTCT	1274
Qy	1201	TCTGCGCTTGGCAAGACAGCTTTAAGCTTCCGAAAGCAATCTCAAGAAAAGTCTCT	1260
Db	1275	TCTGCGCTTGGCAAGACAGCTTTAAGCTTCCGAAAGCAATCTCAAGAAAAGTCTCT	1334
Qy	1261	ACCTCTTCAAGGAGATGATGAGAGACCTTCAAGAACTGCAACCGGTGCTGATATCTCA	1320
Db	1335	ACCTCTTCAAGGAGATGATGAGAGACCTTCAAGAACTGCAACCGGTGCTGATATCTCA	1394
Qy	1321	GAGCAGGCTGTCCCTGCTGATCCCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTCA	1380
Db	1395	GAGCAGGCTGTCCCTGCTGATCCCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTCA	1454
Qy	1381	CAGGTGAATCTGTATCTATCTCTCTCAAGACCCAGAGCTGTGGGAACTTTAGAC	1440
Db	1455	CAGGTGAATCTGTATCTATCTCTCTCAAGACCCAGAGCTGTGGGAACTTTAGAC	1514
Qy	1441	GCCAGTATATGTTTGTGATGAGGAGACCCGCTGAGAGCAAGTCTCCACGCTTCAAGG	1500
Db	1515	GCCAGTATATGTTTGTGATGAGGAGACCCGCTGAGAGCAAGTCTCCACGCTTCAAGG	1574
Qy	1501	CTCCTCTCCACGCAAGGCAATTTCTCAAACTCAATGGCAAGTTCTCCGCAACGCTTA	1560
Db	1575	CTCCTCTCCACGCAAGGCAATTTCTCAAACTCAATGGCAAGTTCTCCGCAACGCTTA	1634
Qy	1561	GAAAGCACTAACCCCTAGCACTTTGGCTCCCTGAGCACTGACCTTCCATCTGCA	1620
Db	1635	GAAAGCACTAACCCCTAGCACTTTGGCTCCCTGAGCACTGACCTTCCATCTGCA	1694
Qy	1621	GCCCGGCGGAGCCGCTTCAAGGGGCTGTATGAGAGCAAGATCTGTCTCCGAGTCC	1680
Db	1695	GCCCGGCGGAGCCGCTTCAAGGGGCTGTATGAGAGCAAGATCTGTCTCCGAGTCC	1754
Qy	1681	TTTGAACCAATTGACCTTGTGCAAGCTTTCCGGAACCCCACTGAGGGGCTGTGTCT	1740
Db	1755	TTTGAACCAATTGACCTTGTGCAAGCTTTCCGGAACCCCACTGAGGGGCTGTGTCT	1814
Qy	1741	GTGAGCAACCTGAGGGGCTTGAAGACCTTCCCTCAAGAGTGTGAAGCATGTGGCAG	1800
Db	1815	GTGAGCAACCTGAGGGGCTTGAAGACCTTCCCTCAAGAGTGTGAAGCATGTGGCAG	1874
Qy	1801	GAAATCTTGGGGATATGCTCTTTTCTGTACAGACTGCCAAGAGTGTGACCTTAC	1860
Db	1875	GAAATCTTGGGGATATGCTCTTTTCTGTACAGACTGCCAAGAGTGTGACCTTAC	1934
Qy	1861	AGAACAAGCCTTAGAATCTGCTCAAGCTAGC	1893
Db	1935	AGAACAAGCCTTAGAATCTGCTCAAGCTAGC	1967

RESULT 4
US-10-343-514-27

; Sequence 27, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEEFEBRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DBA-DRUC/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 2026
; TYPE: DNA
; ORGANISM: RAT
US-10-343-514-27

Query Match 88.9%; Score 1683.4; DB 18; Length 2026;

Best Local Similarity 93.4%; Pred. No. 0;

Matches 176; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 1 ATGAGTGGTGGCTTACTTCCAGCGCCGAGCGAGGCTCCCTGGGCTCCGCGCTGACC 60
DB 75 ATGAGTGGTGGCTTACTTCCAGCGCGGAGGAACTGGGCTCCCTGGGCTCCGCGCTGACC 134
QY 61 TCGGAGGCGCCGCGCGCTGGCGGAGCGGCTCATCAAGTGGCCTTAACTCTGATGAG 120
DB 135 ACGGAGGCGCGCGCGCTGGCGGAGCGGCTCATCAAGTGGCCTTAACTCTGATGAG 194
QY 121 AAGCAGCGGCTGAGAGCGGACCATCAACAACAACCTGCGGACCGCTAGAGTTCTG 180
DB 195 AAGCAGCGGCTGAGAGCGGACCATCAACAACAACCTGAGGACCGCTAGAGTTCTG 254
QY 181 GAGAGCGTGGGCAAGGCGACCTTACGGGAGGTGAAGAGGACGAGAGAGCTCGGGGCT 240
DB 255 GAGAGCGTGGGCAAGGCGACCTTACGGGAGGTGAAGAGGACGAGAGAGCTCGGGGCT 314
QY 241 CTGTGGCCATCAAGTCTCATGAGAAAGCAAAATCAAAATGAGAGAGATCTGTCAC 300
DB 315 CTGTGGCCATCAAGTCTCATGAGAAAGCAAAATCAAAATGAGAGAGATCTGTCAC 374
QY 301 ATACGGAGGAGATGAGATGATCTTCACTCAACCAACCCCAATCATTTGCCATCAT 360
DB 375 ATACGGAGGAGATGAGATGATCTTCACTCAACCAACCCCAATCATTTGCCATCAT 434
QY 361 GAAAGTTTGAATAGCAGCAAGATTGTGATGTCAAGTATGAGATGAGAGAGAGAGAT 420
DB 435 GAAAGTTTGAATAGCAGCAAGATTGTGATGTCAAGTATGAGATGAGAGAGAGAT 494
QY 421 CTGTATATTAATCATGAGTGAAGCGGCAAGCTGAGTGAAGCGGACCGCAGGCAATTTCTTC 480
DB 495 CTGTATACATTAATCATGAGTGAAGCGGCAAGCTGAGTGAAGCGGACCGCAGGCAATTTCTTC 554
QY 481 CGACAGATCGTGTGCGCTGCACTAATGCGCAAGAGCGGAGCGTTCACCGAGATGTC 540
DB 555 CGACAGATCGTGTGCGCTGCACTAATGCGCAAGAGCGGAGCGTTCACCGAGATGTC 614
QY 541 AAGCTGAAAAATCCTTCTAGATGCAATGAAATCAATCAAGTTGCTGATTTGGCTC 600
DB 615 AAGCTGAAAAATCCTTCTAGATGCAATGAAATCAATCAAGTTGCTGATTTGGCTC 674
QY 601 TCCAACTGTATCAAAAGGCAAGTTCTTCCAGAGCTTCTGAGAGCGCTCTCTAGGCC 660
DB 675 TCCAACTGTATCAAAAGGCAAGTTCTTCCAGAGCTTCTGAGAGCGCTCTCTAGGCC 734

QY 661 TCGCTGAGATGATCAACGGGAGCCCTTATGTGGGCCCAAGGTGACAGCTGTCTG 720
DB 735 TCACCTGAGATCGCAACGGGAGCCCTTATGTGGGCCCAAGGTGACAGCTGTCTG 794
QY 721 GGGTCTCTGTATCACTCTGTGATGAGGACCAATGCGCTTTGACGGGACAGATCATATA 780
DB 795 GGGTCTCTGTATCACTCTGTGATGAGGACCAATGCGCTTTGACGGGACAGATCATATA 854
QY 781 ACACTGTGAAAGCAATCAATGAGGAGCTTACCGTAGCGGACCGCAAGCCCTCCAGATCC 840
DB 855 ACCCTGTGAAAGCAATCAATGAGGAGCTTACCGTAGCGGACCGCAAGCCCTCCAGATCC 914
QY 841 TGTGGCTGATCGGCTGCTGTTAATGATGAAACCCCAACCGTGGGCTCACACTGAGAT 900
DB 915 TGTGGCTGATCGGCTGCTGTTAATGATGAAACCCCAACCGTGGGCTCACACTGAGAT 974
QY 901 GTAGCCAGTCAATGATGATGATCACTGGGCTTACCAACCGAGTCCGGGAGACAGAAACC 960
DB 975 GTAGCCAGTCAATGATGATGATCACTGGGCTTACCAACCGAGTCCGGGAGACAGAAACC 1034
QY 961 CTGGTGAAGGATGAGGACCCCTAGTGTGATCTTGGCGGCTCCATGAGCGGATGATTA 1020
DB 1035 CTGGTGAAGGATGAGGACCCCTAGTGTGATCTTGGCGGCTCCATGAGCGGATGATTA 1094
QY 1021 CGTGGCTCTTGGCGGCTCCCTCTGAGAAATGAGACCAAGTGTGCAAGCTTCTTCAAGCAG 1080
DB 1095 CGTGGCTCTTGGCGGCTCCCTCTGAGAAATGAGACCAAGTGTGATGATCTTCTTCAAGCAG 1154
QY 1081 CAGCTGCCGAGGATGAGACATGATCTGAGGCTGAGAGCGGCAACATTTCTTAAAGAG 1140
DB 1155 CAGCTGCCGAGGATGAGACATGATCTGAGGCTGAGAGCGGCAACATTTCTTAAAGAG 1214
QY 1141 TCCGAAAGGAGATGACATGAGCTCAAAATCTGCAAGGTGACCGGCTGAGATACCTCT 1200
DB 1215 TCCGAAAGGAGATGACATGAGCTCAAAATCTGCAAGGTGACCGGCTGAGATACCTCT 1274
QY 1201 TCTCGCCCTGCAAGACAGCTTAACTTCCGAAAGGCAATTTCAAGAAATGCTCTCT 1260
DB 1275 TCTCGCCCTGCAAGACAGCTTAACTTCCGAAAGGCAATTTCAAGAAATGCTCTCT 1334
QY 1261 ACTCTGCAAGGAGATGACAGAGACCTTCAAGAACTCAGACCGGTGCTGATATCTCA 1320
DB 1335 CCTCATCTGGGAGGATGACAGAGGCGCTTCAAGAACTCAGACCGGTGCTGATATCTCA 1394
QY 1321 GGGCAGCTGTCTCTGTATCTCTCTCCAGAAAGGCAATTTTAAAGATCTGCA 1380
DB 1395 GGGCAGCTGTCTCTGTATCTCTCTCCAGAAAGGCAATTTTAAAGATCTGCG 1454
QY 1381 CAGCGTAATCTGTATCTACTCTCTCCAGAGCCAGCGATCTGGGAACTCTTACAG 1440
DB 1455 CAGCGTAATCTGTATCTACTCTCTCCAGAGCCAGCGATCTGGGAACTCTTACAG 1514
QY 1441 GCCAGTATGTTTGTGAGTGGGACCCCGTGAAGAGAGTCTTCAAGAGCTTCAAGG 1500
DB 1515 GCAAGTATGTTTGTGAGTGGGACCCCGTGAAGAGAGTCTTCAAGAGCTTCAAGG 1574
QY 1501 CTCTCTCTCAACCGGAGGATCTCAAACTCAATGAGAGTCTCCCGACAGCTTCA 1560
DB 1575 ---CGCTCTCAATGCAAGGAGATCTCAAACTCAATGAGAGTCTCCCGACAGCTTCA 1631
QY 1561 GAAAGCACTACCCCTTACAGCTTGGCTCCCTGAGCAACTGAGCTCTCCATCTGCA 1620
DB 1632 GAAAGCACTACCCCTTACAGCTTGGCTCCCTGAGCAACTGAGCTCTCCATCTGCA 1691
QY 1621 GCCGAGCCAGAGCGGCTTCAAGGAGGCTGAGAGAGAGATCTGTCTCCGAGTCC 1680
DB 1692 GCCGAGCCAGAGCGGCTTCAAGGAGGCTGAGAGAGAGATCTGTCTCCGAGTCC 1751
QY 1681 TTTGACCAATTTGACCTTGGCTGAAAGCTTCCGAAACCCCACTGAGAGGCTGTGTCT 1740
DB 1752 TTTGACCAATTTGACCTTGGCTGAAAGCTTCCGAAACCCCACTGAGAGGCTGTGTCT 1811

QY	1741	GTGGAACAACCTGAGAGGGGCTTGACACGCTCCCTCAGAGAGTCTGAAGCCATGTGTGGCAG	1800
Db	1812	GTGGAACAACCTGAGAGAGGCTTGAACACGCTCCCTCAGAGAGCTTAAACGATGTGTGGAG	1871
QY	1801	GAATCTTGTGGGGGATAGAGCTGCTTTTCTCTGACAGCTGGCCAAAGAGTGAATCTGACAGCCTAC	1860
Db	1872	GAATCTTGTGGGGGATAGAGCCTTTTCTCTGACAGCTGGCCAAAGAGTGAACAGAGCCTTAC	1931
QY	1861	AGACAAAGCCCTAGAGATCTGCTCAAAAGCTCAGC	1893
Db	1932	AGACAAAGCCCTAGAGATCTGCTCAAGCTCAGC	1964

RESULT 5

```

US-10-343-514-1
? Sequence 1, Application US/10343514
? Publication No. US20040132025A1
? GENERAL INFORMATION:
? APPLICANT: DRUCKER, Daniel J.
? APPLICANT: ROSEN, Cheryl F.
? APPLICANT: LEBEVRRE, Diana L.
? TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
? FILE REFERENCE: DPA-DRUC2/PCT
? CURRENT APPLICATION NUMBER: US/10/343,514
? CURRENT FILING DATE: 2003-01-31
? PRIOR APPLICATION NUMBER: PCT/CA01/01109
? PRIOR FILING DATE: 2001-08-02
? PRIOR APPLICATION NUMBER: US 60/222,650
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: US 60/274,613
? PRIOR FILING DATE: 2001-03-12
? PRIOR APPLICATION NUMBER: CA 2,340,780
? PRIOR FILING DATE: 2001-03-28
? NUMBER OF SEQ ID NOS: 109
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 1
? LENGTH: 2929
? TYPE: DNA
? ORGANISM: RAT
? US-10-343-514-1

```

Query Match	88.9%	Score 1683.4;	DB 18;	Length 2929;
Best Local Similarity	93.4%;	Pred. No. 0;		
Matches 1769;	Conservative	0;	Mismatches 121;	Indels 3;
			Gaps	1;

[illegible]

Qy	421	CTGTATATATTAACATCAATGAGCGGGCCACGGCTAGATGAGCGGGACCGCACAGGCAATTTCTTC	480
Db	503	CTGTATCAATTAACATCAATGAGCGGGCCACGGCTGATATAGCGGGACCGCACAGGCAATTTCTTC	562
Qy	481	CGACAGATTCGTGTCCCTGCACATACTGCACACAGAAAGGGATCGTTTCAACGAGATCTC	540
Db	563	CGACAGATTCGTGTCCGCCCTGCACATACTGCACACAGAACGGGATGTTTTCACCGGGACCTC	622
Qy	541	AACTGGAAAAATCTCTTCTAGATGCGCAATGAAAACATCAAGATTGCTGACTTTGGCTTC	600
Db	623	AACTGGAGAAACATCTCTTAGATGCGCAGTGGGCAACATCAAGATTGCTGATTTGGCTTC	682
Qy	601	TTCAAACCTGTATCCACAAAGGCAAGTTCTCTCAACGTTCTGTGGGACCTCTCTTAAGCC	660
Db	683	TTCAAACCTGTATCAAAAGGCAAGTTCTCTCAACGTTCTGTGGGACCTCTCTTAAGCC	742
Qy	661	TCGCTCAGATATGTCACACGGGAAGCCCTATGTGGAGCCACAGAGGTGACAGCTGTCTCTG	720
Db	743	TCACCTTAGATCGTCAACGGGAAGCCCTATGTGGAGCCACAGAGGTGACAGCTGTCTCTG	802
Qy	721	GAGCGTCTCTGTATCATCTGTGTGCATGGACCATATGCCCTTTGACGGGACAGATCATAAA	780
Db	803	GAGCGTCTTCTGTATCATCTGTGTGCATGGACCATATGCCCTTTGACGGGACAGATCATAAA	862
Qy	781	ACACTGTGAAAGCAAAATCATGTAACGGGGCTTTACGTATGCGGCCACAGCCGCTCGATGCC	840
Db	863	ACCTGTGTGAAACAAATCAATGAGGGGGCTTACCGAAGGCCGTGTGATGATGCC	922
Qy	841	TGTGGCCTGATCCGGTGGGCTGTTAAATGTGTGACCCACCCGTCGGGGCACATGGAAGAT	900
Db	923	TGTGGCCTGATCCGGTGGGCTGTTAAATGTGTGATCCATCCGTGGGGCACCTTGAAGAT	982
Qy	901	GTAGCCAGTATTGTGTGGTCAACTGGGGTTTACACACACCGAGTCGGGGAAACAGAAAGCC	960
Db	983	GTAGCCAGTATTGTGTGGTCAACTGGGGTTTACACACCCGAATTGGGGAAACAGAAAGCT	1042
Qy	961	CTGCGTAAAGGTGGGACCCCTAATGTGTGACTTTTGCCGGGCTTCATAGCGGAACTGTGTTA	1020
Db	1043	CTGCGAAGAGGTGGGACCCCTAAGGGTGTACTTGTGGCGGGCTCTATATGGGGAATGTGTTA	1102
Qy	1021	CGTGGCTCTCGAGGCCCTCTCTGTGAAGATGTGAGCCAAAGGTGTGACGCTTCTTAAAGAG	1080
Db	1103	CGTGGCTCTCTCGAGGCCCTCTCTGTGAAGATGTGAGCCAAAGTGTGTACTTCTTAAAGAG	1162
Qy	1081	CACGTGCGGGAAGGTGGAACACTGTACTGTGGCTGTGAGCGGCACAACTTCTTTAAGAG	1140
Db	1163	CATGTGCGGGAAGGTGGAACAACGGGACCGGGGCTGGAAGGGCAACTTCTTTAABAG	1222
Qy	1141	TCCGGAAGGAATGACATATGCTCAAAATCTGTCAAGGTACCCGGCTGAGATACCTCT	1200
Db	1223	TCCCGCAAGGAATGACATATGCTCAAACTCTGTAGATGACCCAGTTGAAGATCTTCC	1282
Qy	1201	TCTGCGCTGCGCAAGACACCTTAAGCTTCCGAAAGGCAATTCAGAAAAAGTCCCTCT	1260
Db	1283	TCTGCGCTTCCGCAAGAACACCTCTCAACCTTCCGAAAGTATCTCAAGAAAAAGGCTCT	1342
Qy	1261	ACCTCGTCAAGGGAGGTACAGAGAGACCTCAGGAATCAGACCGGTGCTGTATCTCA	1320
Db	1343	CCCTCATCGGGGGAAGGTACAGAGAGGCGCTCAGGAATCAGACAGTGTCCATATCCCA	1402
Qy	1321	GAGGACGCTGTCCCTGTATCTATCTCTGTCTGCCAGAGAAAGGCATCTTTAAGAACTCTGA	1380
Db	1403	GAGGACGCTGTCCCTGTATCTATCTCTGTCTGCCAGAGCCAGTATGTGGGAACTCTTAGAC	1462
Qy	1381	CAGGTGAAATCTGTTATCTATCTCTCTCCAGAGCCGACAGTCTGGGGAACTTTTAGAC	1440
Db	1463	CAGGTGAAATCTGTTATCTATCTCTCTCCAGAGCCAGTATGTGGGAACTTTTAGAC	1522
Qy	1441	GCCAGTATATGTTTGTGAGTGGGGACCCCGTGGAGCAAGTCTCCACAGGCTTACAGG	1500
Db	1523	GCAAGTATATGTTTGTGAGTGGGGACCCCGTGGAGCAAGTCTCCACAGGCTTACAGG	1582
Qy	1501	CTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGAAGTTCTCCGCAACGCTTAA	1560

1583 ---CGCTTCATGCAAGGAGATCTCAAACTCAATGCAAGTTTCCCGACAGCTTTA 1639
1561 GAAGGCACTACCCCTAGACACTTTGGCTCCCTGGAGCAACAGTGGCTCCTCCATCTGCA 1620
1640 GAAGGCACTGCTCCCTAGACACTTTGGCTCCCTAGACCAACAGTGGCTCCTCCATCTGCA 1639
1621 GCCCGGCGCAAGCGCCCTGAGGGGCTGTAGAGAGACAGCACTGTCTCTCCAGTCC 1680
1700 GCCCGGCGCAAGCGCTCCTGAGGGGCTGTAGAGAGACAGCACTGTCTCTCCAGTCC 1759
1681 TTTGACCAATTTGACCTTGGCTGAACTGTCTCCGAAACCCCACTGAGGGGCTGTGTCT 1740
1760 TTTGACCAATTTGACCTTGGCTGAACTGTCTCCGAAACCCCACTGAGGGGCTGTGTCT 1819
1741 GTGACCAACCTGAGGGGGCTTGGAGCAGCTCCCTCAGAAAGGTGAGAGATGTGTGAG 1800
1820 GTGACCAACCTGAGAGAGGCTTGGAGCAGCTCCCTCAGAAAGGTGAGAGATGTGTGAG 1819
1801 GAATCTTGGGGGAGTACCTGCTTTCTGTGACAGACTGCGAAGAGTGTGAGCTGAC 1860
1880 GAATCTTGGGGGAGTACCTGCTTTCTGTGACAGACTGCGAAGAGTGTGAGCTGAC 1939
1861 AGACAAAGCCTTAGAATCTGTCTCAAGCTCAGC 1893
1940 AGACAAAGCCTTAGAATCTGTCTCTCAAGCTCAGC 1972

RESULT 6

US-09-963-159-3
; Sequence 3, Application US/09963159
; Patent No. US2002007312A1
; GENERAL INFORMATION:
; APPLICANT: Curtin, Rory A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-5001
; CURRENT APPLICATION NUMBER: US/09/963,159
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-159-3

Query Match 71.9%; Score 1360.2; DB 9; Length 1884;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 156; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

QY 47 CCTCCGCGCTGAGCGCCGCGCTGCGGAGCGGGCTCATCAAGTCGCTTA 106
DB 35 CCACTCCCTCGGCGGAGAGCTAGCCGCGCGCTGCGGAGGGCTGATCAAGTCGCTCA 94
QY 107 AACCTCTGATGAAGAGAGAGCGGTGAAGGGGCACTCAAAACAACAACCTTGGGCAAC 166
DB 95 AGCCCTTAATGAAGAGAGAGCGGTGAAGGGGCACTCAAAACAACAACCTTGGGCAAC 154
QY 167 GCTACGAGTCTCTGAGAGCGCTGAGAGGCACTCAAGGAGAGTGAAGAGGCAAG 226
DB 155 GCTACGAGTCTCTGAGAGCGCTGAGAGGCACTCAAGGAGAGTGAAGAGGCGG 214
QY 227 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCATCAGAAAGCAAAATCAAGATGAGC 286
DB 215 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCATCAGAAAGCAAAATCAAGATGAGC 274
QY 287 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCATCAGAAAGCAAAATCAAGATGAGC 346
DB 275 AAGATCTGATGACATACGAGGAGATGATGATCATGTCATCACTCAACCACTCTACA 334

QY 347 TCATTCGATTCATGAAGTGTGAGATATGACAGCAAGATTGTGATGTGATGAGATG 406
DB 335 TCATTCGATTCATGAAGTGTGAGATATGACAGCAAGATTGTGATGTGATGAGATG 394
QY 407 CCAGCGAGGCGATCTGTATGATTAATCATCATGTAGGGGCGCAGGCTGAGTGGGGGAG 466
DB 395 CCAGCGAGGCGATCTGTATGATTAATCATCATGTAGGGGCGCAGGCTGAGTGGGGGAG 454
QY 467 CCAAGCATTTCTTCCGACAGATCTGTCTGCGCTGACATCTGACACAGAACGGATGCG 526
DB 455 CTAGGATTTCTTCCGACAGATCTGTCTGCGCTGACATCTGACATGAGAACAGATG 514
QY 527 TTACCGAGATCTCAGAGTGAAGAAATCTTCTGATATGCCATGAAACATCAAGATTG 586
DB 515 TCACCGAGATCTCAGAGTGAAGAAATCTTCTGATATGCCATGAAATATCAAGATTG 574
QY 587 CTGACCTTGGCTCTCAACCTGATCAACAAAGGCAAGTCTCTCAGACGTTCTGTGGGA 646
DB 575 CTGACCTTGGCTCTCAACCTGATCAACAAAGGCAAGTCTCTCAGACATTCGTGGGA 634
QY 647 GCCCTCTCTACGCGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCCAGAGTGG 706
DB 635 GCCCTCTCTACGCGCTGAGATGATCAACGGGAAAGCCCTATGAGGAGCCAGAGTGG 694
QY 707 ACAAGTGTCTCTGGGCGTTCTCTGTATCTCTGTATCTGTGATGAGCAATGCTTTGAG 766
DB 695 ACAAGTGTCTCTGGGCGTTCTCTGTATCTCTGTATCTGTGATGAGCAATGCTTTGAG 754
QY 767 GGCAGGATCATTAACACTGTGAGCAATCAATCAAGGGGCTTACCTGTAGGCGGCCA 826
DB 755 GGCAGGATCATTAACACTGTGAGCAATCAATCAAGGGGCTTACCTGTAGGCGGCCA 814
QY 827 AGCGGTGATGCTGAGGCTGTATCGGAGGCTGTAAATGATGAGCAAGCCGCTGGG 886
DB 815 AACCTTGAATGCTGAGGCTGTATCGGAGGCTGTAAATGATGAGTAAACCCAGCGGG 874
QY 887 CCAACTGAGAGATGATGAGCAAGTATGATGAGTCACTGGGGTTAACAACCGAGAGTGG 946
DB 875 CCAACTGAGAGATGATGAGCAAGTATGATGAGTCACTGGGGTTAACAACCGAGAGTGG 934
QY 947 GGGAAACAGAAAGCTGAGAGGAGTGGGCACTTATGATGATGATGATGATGATGATG 1006
DB 935 GAGAGCAAGAGGCTGAGAGGAGTGGGCACTTATGATGATGATGATGATGATGATG 994
QY 1007 TGGGGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 995 TGGGGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1054
QY 1067 GCTTCTTCAAGAGCAAGTGTGAGAGGAGTGAAGCACTGAGCTGAGAGCGGCAAC 1126
DB 1055 GCTTCTTCAAGAGCAAGTGTGAGAGGAGTGAAGCAACCTGAGCTGAGAGCGGCAAC 1114
QY 1127 ATTCTCTTAAGAGTCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1186
DB 1115 ATTCTCTTAAGAGTCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1174
QY 1187 CTGAGATTAATCTTCTGAGGAGTGAAGAGGCTTAAAGCTTCCGAAAGGCAATTC 1246
DB 1175 CTGAGATTAATCTTCTGAGGAGTGAAGAGGCTTAAAGCTTCCGAAAGGCAATTC 1234
QY 1247 AGAAAGATCTCTACTCTGTCAGAGGAGTGAAGAGGAGTGAAGAGGAGTGAAGAGG 1306
DB 1235 AGAAAGATCTCTACTCTGTCAGAGGAGTGAAGAGGAGTGAAGAGGAGTGAAGAGG 1294
QY 1307 TGCCTGATCTCAGAGGAGCTGTCTCTGTATCTCTGTCTTCCAGAGAAAGCAATCC 1366
DB 1295 TGCCTGATCTCAGAGGAGCTGTCTCTGTATCTCTGTCTTCCAGAGAAAGCAATCC 1345
QY 1367 TTAAGAGTCTGAGAGGAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATG 1426
DB 1346 TTAAGAGTCTGAGAGGAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATG 1405
QY 1427 GGGAACTCTTAAGAGGAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATG 1486

QY	887	CCACACTGGAGATGTATGACAGTCAATGGTGGGTGAACATGGGGTTTAAACACACGGAGTGC	946
Db	875	CCACCTTGGAGGATGTGGCCACTCACTGGTGGGTGAATGGGGCTTACGCCACCCGATGG	934
QY	947	GGGAAACAGAAACCCCTGCGTGAAGGGTGGCACTTAATGTATCACTTTGGCCGGGCTTCA	1006
Db	935	GAGACAGAGAGGCTCGCATGAGGGTGGCACTTGGCAGTGAATCTGTGCCCGCGCTTCA	994
QY	1007	TGGCCGACTGGTTAAGTGGCTTCTTGGCGCCCTCTTGGAGAAATGAGATCGAAGTGTGCA	1066
Db	995	TGGCTGACTGGCTCGGCGTTTCTCCCGCCCTCTTGAAGAAATGGGGCAAGGTGTGCA	105
QY	1067	GCTTCTTCAAGCAGCAGTGGCCGGGAGGTTGAAGACATGTACTCTGGGCTGGAGCGGCAAC	1122
Db	1055	GCTTCTTCAAGCAGATGCACTGGTGGGGAGAACACACCTTGGCTGGAGCGCGAC	111
QY	1127	ATTCTTAAAGAAAGTCCGGAAAGAAATGACATGGCTCAAAATCTGCAGGTGAACCGG	1186
Db	1115	ATTGCTCAAGAAATGCCGCAAGGAATGACATGGCCAGTCTCCACAGTGAACGG	117
QY	1187	CTGAGGATTAACCTTCTTGGCCCTTGGCAAGACGCTTAAGTTTCCGAAAGGCAATTTCA	1244
Db	1175	CTGAATGACATGCCCATCGCTCTTGGCAAGAACCTTAAGTTTCCGAAAGGCAATTTCA	123
QY	1247	AGAAAAGTCTCTACCTTGTGAGGGGAGTTACAGAGGACCTCAGGAACTCAGACGG	1306
Db	1235	AGAAAGATGTCAACCTCTTGAAGAAAGGGTACAGAGGACCTTCGAGTCAACCCAA	129
QY	1307	TGCTGATATCTCAGGGGACCTGTCTCTGTATCTCTGCTGCCAAGAAAGGCATCC	1366
Db	1295	TCCCTCGAGCCCAAGGGCAGGCTGCCCC-----CCGTGCTCCCAAGAGGCATTC	1345
QY	1367	TTAAGATCTGACAGCGTGAATCTGATTACTCTCTCTCAGAGCCAGCGAGTCTG	1422
Db	1346	TCAAGAGCCCGACAGCGAGTCTGGCTACTACTCTCTCCAGGCCAATGAAATCTG	1405
QY	1427	GGGAATCTTGAACCCAGTAGTATGTTTGTGATGAGGGGACCCCGTGAAGCAGAATCTC	1486
Db	1406	GGGAACTCTTGAACCCAGCGCAGCGATGTTTGTGAATGGGGAATCCAAAGAGCAGAACTTC	1465
QY	1487	CACAGGCTTCAAGGGCTCTCTCTCAACCGCAGAGGCAATTTCAATCAATGGCAATTT	1546
Db	1466	CGCAAGCTTCAAGGGCTCTCTCTCAATCGCAAGAGCATCTCAAACTCAATGGCAATTT	1522
QY	1547	CCCGACAGCCCTTAAGGCACTAACCCCTAGACCTTTTGGCTCCCTGAGCCAACTGGCT	1608
Db	1526	CCAGACAGCCCTTAAGGCTCGGGGCCCAACCACTTGGGCTCTCTGATGAATCTGGCC	1588
QY	1607	CTTCCATCTTGAAGCCCGGCCAGCCGCTCTCAAGGGCTGTGATGAGCAACATTC	1666
Db	1586	CACCTCGCCCCCTGGCCGGGGCAACCGACCTTCAAGGGCTGTGAGCGAGCAACATTC	1645
QY	1667	TGTCTCTCGAGTCTTTTGAACAAATTGGAATTGCTGAAAGCTTTTCCGAAACCCCACTGA	1722
Db	1646	TGTCTCTCGAGTCTTTTGAACCAAGCTGGAATTCCTGAACCGCTCCCAAGCCCCCACTGC	1705
QY	1727	GGGGCTGTGTGTCTGTGGAACAACCTGAAGGGGCTTGAAGCAGCTCCCTCAGAAAG-----	1780
Db	1706	GGGGCTGTGTGTCTGTGGAACAACCTCAAGGGGCTTGAAGGCCCCCTTCAAGGGGCTTG	1765
QY	1781	-----GTCTGAAGCAATGTGTGGCAAGAAATCTTGGGGATGAAGCTTTTCTCTGAACG	1834
Db	1766	GAAAGCTGCTGAGGGCTGGCGGAGAAATCTTTGGGGGAGACGTGCTTTTCTCTGAACG	1822
QY	1835	ACTGCAAGAGGTGACTGACGCTTACAGCAGACGCTTGAAGAAATGTGCTCAAGGCTCAAG	1893
Db	1826	ATTGCAAGAGGTGACGCACTTACCAAGCAGCAGCTGAGGGTGTGCTTCAAGGCTCAACC	1884

RESULT 8
US-09-963-159-1
; Sequence 1, Application US/09963159
; Patent No. US20020077312A1

```

: GENERAL INFORMATION:
: APPLICANT: Curtis, Roy A.J.
: APPLICANT: Galvin, Katherine M.
: TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
: FILE REFERENCE: 10147-5001
: CURRENT APPLICATION NUMBER: US/09/963,159
: CURRENT FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US 60/234,922
: PRIOR FILING DATE: 2000-09-25
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
:
: LENGTH: 3353
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: unsure
: LOCATION: (3268)
: NAME/KEY: unsure
: LOCATION: (3270)
: NAME/KEY: unsure
: LOCATION: (3272)
:
: US-09-963-159-1

```

Query Match	71.9%;	Score 1360.2;	DB 9;	Length 3353;
Best Local Similarity	84.2%;	Pred. No. 0;		
Match 1565; Conservative	0;	Mismatches 273;	Indels 21;	Gaps 2;

QY	47	CCTCCGCGCTGCGCTCGGAGAGCCGCCGCGCTGCGGAGACGGGCTCATCAAGTCGCGTA	106
DB	191	CCATCTCCCTCGCGCGGAGAGCTACCCCGCGCGTGGGAGAAAGGCTCATCAAGTCGCGCA	250
QY	107	AACCTCTGATGAAGACGAGCGGCTGAAGCGGACCATCACAACACACCTCGCGCAC	166
DB	251	AGCCCTTAATGAAGAACAGCGCGGTGAAGCGGCACCAACAAGCAACCTCGCGCAC	310
QY	167	GCTACGAGTTCCTTGAGACACGCTGGGCAAGGCAACCTACGGGAAGGTGAAGAGGCAAG	226
DB	311	GCTACGAGTTCCTTGAGACACCTGGGCAAGGCAACCTACGGGAAGGTGAAGAGGCGCGG	370
QY	227	AGAGCTCGGGGCGTCTGGTGGCCATCAAGTCATCAAGAAACAAATCAAAATGATAGC	286
DB	371	AGAGCTCGGGGCGCTGGTGGCCATCAAGTCATCAAGAAACAAATCAAAATGATAGC	430
QY	287	AGAGTCTGCTGCATACACGAGGAGATTGATCATGTCTTCACTCAACCAACCCCA	346
DB	431	AAGATCTGATGACATACGAGGAGATTGATCATGTATCATCACTCAACCAACCTCA	490
QY	347	TCATTTGCCATCTCAAGAAAGTTTGAAGATAGCAAGATTTGATTTGATGAGATAG	406
DB	491	TCATTTGCCATCTCAAGAAAGTTTGAAGAAAGCAAGATTTGATTTGATGAGATAG	550
QY	407	CCACCCGAGGATCTGATATGATTAATCATGATGAGGGGCAACGCGCTGATGAGCGGGAC	466
DB	551	CCACCCGAGGATCTGATATGATTAATCATGATGAGGGGCAACGCGCTGATGAGGGGAC	610
QY	467	CCAGGCAATTTCTTCCGACAGATCGTGTCTGGCCCTGCATCTAGGCAACGAACGGATCG	526
DB	611	CTAAGGCAATTTCTTCCGCAAGATCTCTCTGCGCGTGCATATTTGGCATCAAGACGAATG	670
QY	527	TTCAACCGAGATCTCAAGCTGAAAAACATCTCTTAGATGCCAATGAAACATCAAGATTG	586
DB	671	TTCAACCGAGATCTCAAGCTGAAAAACATCTCTTAGATGCCAATGAAATATCAAGATTG	730
QY	587	CTGACTTTGGGCTCTCCAACTCTGTAACAACAAGGCAAGTTCTCTCAACGTTCTGTGGGA	646
DB	731	CTGACTTTGGGCTCTCCAACTCTGTAACAACAAGGCAAGTTCTCTGCAACATTTCTGTGGGA	790
QY	647	GGCCTCTCTACGCTCGCTGAGATGATCAACGGGAAGCCCTTATGTGGGCGACAGAGTGG	706
DB	791	GGCCTCTCTACGCTCGCAGAGATTTGATCAATGGGAAGCCCTTACACAGGCCACAGAGTGG	850
QY	707	ACAGTGGTCTCTGGGCGGTTCTCTGTATCATCTGTGATCAAGGACCATGACCTTTGAC	766

QY	107	AACCTGTAATGAAGACGAGCGGTGAAGCGGCAACATCAAAACAACAACTGGCGCAC	166
Db	251	AGCCCTTATATAGAAGACGAGCGGTGAAGCGGCAACATCAAAACAACAACTGGCGCAC	310
QY	167	GCTACGAATTCCTGGAGACGCTGGGCAAGGGCACTTACGGGAAGGTGAAGAAGGCACGAG	226
Db	311	GCTACGAATTCCTGGAGACCTTGGGCAAAAGGCACCTTACGGGAAGGTGAAGAAGGCAGG	370
QY	227	AGAGCTCGGGCGCTGTGGTGCATCAAGTCCATCAGGAAGAACAATCAAAATGAGC	286
Db	371	AGAGCTCGGGCGCGCTGTGGTGCATCAAGTCAATCCGAGAGAACAAATCAAAATGAGC	430
QY	287	AGAGTCTGCTCAACAACGAGAGGAGATTTGAGTCAATCTTCACTCAACCAACCCCA	346
Db	431	AAGATCTGATCAACAACGAGAGGAGATTTGAGTCAATCTTCACTCAACCAACCCCA	490
QY	347	TCATTTGCATCATGAAGTGTGTTGAGATTAACAGCAAAATGTGATTTGTCAATGAGATG	406
Db	491	TCATTTGCATCATGAAGTGTGTTGAGAACAGCAGCAAAATGTGATTTGTCAATGAGATG	550
QY	407	CCAGCCGAGCGGATCTGTATGATTAACATCAATGAGCGGCAACGCTGAGTGGCGGACG	466
Db	551	CCAGCCGAGCGGACCTTATATGATTAACATCAAGGAGCGGCAACGCTGAGTGGCGGACG	610
QY	467	CCAGGCATTTCTTCCGACAGATCGTGTCTGCCCTGCACTAATGCCACAGAACGGATGCG	526
Db	611	CTAGGCATTTCTTCCGAGAGATCGTGTCTGCCGTGCACTAATTTGCATCAGAACAGATTTG	670
QY	527	TTCAACCGAGATCTCAAGCTGGAAAAACATCCTTTAGATGGCAATGGAAACATCAAGATG	586
Db	671	TTCAACCGAGATCTCAAGCTGGAAAAACATCCTTTAGATGGCAATGGAAATATCAAGATG	730
QY	587	CTGACTTTGGGCTCTCCAACCTGTATCACAAAAGGAGCAATTCCTCAGACGTTCTGTGGGA	646
Db	731	CTGACTTGGGCTCTCCAACCTGTATCACAAGGAGCAATTCCTGTGAGCAATTTGTGGGA	790
QY	647	GCCCTCTTACGCTCGCTGTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCAAGATGG	706
Db	791	GCCCCCTTATGCTCTCGCCAGAGATTTGCAATGGGAAGCCCTTACACAGGCCCAAGATGG	850
QY	707	ACAGCTGGTCTCTGGGAGGTCTCCGTGATCACTCTGGTGCATAGGGCAACAATGCCCTTTGACG	766
Db	851	ACAGCTGGTCTCTGGGAGGTCTCTCTCTATCAATCTGTGTCAATGGAACAACATGCCCTTTGATG	910
QY	767	GGCAGGATCATTAACAATGTGTGAAGCAAAATCAGTAAAGGGGCTTTACCGTGAAGCGCCCA	826
Db	911	GGCATGACATTAAGATCTCTAGTGAACAAGATCAGCAAGGGGCTTACCGGGAGGCACCTA	970
QY	827	AGCGGTCCGATGCTGTGGCTGTATCCGGTGGCTGTTATATGTGTGAACCCACCCGTGGG	886
Db	971	AACCTCTGATGCGCTGTGGCTGTATCCGGTGGCTGTTATATGTGTGAACCCACCCGCGGG	1030
QY	887	CCAACCTGGAAGATGTACCAATGATTTGTGTGGTCAATCTGGGGTTACACACCGGAGTGC	946
Db	1031	CCACCCTGGAAGATGTGTGCAATGATTTGTGTGGTCAATCTGGGGTTACACCAACCGGAGTGC	1090
QY	947	GAGAAACAGGAAGCCCTGCGTGAAGGGTGGGCAACCTTATGATGTGACTTTTGGCGGGCTTCA	1006
Db	1091	GAGAGCAGGAAGCTTCCGATGAGGGTGGGCAACCTTGGAGATGATCTGTGCCCGGCTTCA	1156
QY	1007	TGGCGGACTGTGTATGTCGCTTCTCGCGGCCCTTCTGTGAAGATGGAACCAAGATGTGCA	1066
Db	1151	TGGCGGACTGTGTATGTCGCGGCTTCTCGCGGCCCTTCTGTGAAGATGGAAGCAAGATGTGCA	1210
QY	1067	GCTTCTTCAACAGACGATGCCCGGGAAGGTGAAGACATGTACCTGGGTGTGAAGCGGCAAC	1126
Db	1211	GCTTCTTCAACAGACGATCACTGTGTGGGAAGCAACCCCTGTGGCTGTGAAGCGGCAAC	1270
QY	1127	ATTCTCTTAAAGATGCCGAAAGAGATGACATGCTCAAAATCTGCAAAGTGAACCGG	1186
Db	1271	ATTGCGCTCAAGAAATGCCGAAAGAGATGACATGCGCCAGTCTTCCACAGTGAACCGG	1330
QY	1187	CTGAGGATACCTCTTCTCGCCCTGGCAAGAGCAAGCTTAAAGCTTCCGAAGGCAATTTCA	1246

Db	1331	CTGATGACACTGCGCCATTCGCCCTTGGCAAGACCACTTCAGCTCCCAAGGGCAATTCTCA	1390
Qy	1247	AGAAAAAGTCCTTAACCTTCGTCAAGGAGGTACAGAGACCTTCAGAACTCAGACCGG	1306
Db	1391	AGAAAGAGGTGACACCTCTTGCAAGAGGGGTACAGAGAGACCTCCCGAGCTCAGCCCAA	1450
Qy	1307	TGCCTGATATCTCCAGGGGACGCTCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCATCC	1366
Db	1451	TCCCTCGACGCCCAGGGGACGGCTGCC-----CCCTGCTCCCAAGAGGGCAATTC	1501
Qy	1367	TTAAGAGTTCGACAGCGTGAATCTGGTATCTACTCCTCTCCAGAGCCCGACGACTCG	1426
Db	1502	TCAAGAACCCCGACAGCGGAGGTCTGGCTATCTACTCTCTCCCGAGCCGATGAATCTG	1551
Qy	1427	GGGAATCTTTAGACGCCAGTATGTGTTGTAGATGGGAGCCCGTGGAGCAGAATCTC	1486
Db	1562	GGGAGCTTTGAGACGACAGGCGACGTGTTGTGATGGGGATCCCAAGAGCAGAAAGCTC	1621
Qy	1487	CACAGGCTTCAGGGGCTCTCTCCACCGCAAGGGCAATTCGAACTGAATGGCAATTC	1546
Db	1622	CGAAGCTTCAGGGGCTCTCTCCATCGCAAGGCATCTCMAACTCAATGGCAAGTCT	1681
Qy	1547	CCCGACAGCCTTGAAGAGCAATACCCCTGACACTTGGCTCCCTGGACCAATGGGCT	1606
Db	1682	CCGACACGCTTGAAGCTCGGGCCCCCACACTTCGCTCCCTGATGAATCTCGCC	1741
Qy	1607	CCTCCCATCTCGACGCCCGGCCACGCCCTTCAGGGGCTGTAGTGAAGACAGCATCC	1666
Db	1742	CACCTCGCCCTTGCGCCCGGCCACGCCACCTCAGGGGCTGTAGAGCAGACATCC	1801
Qy	1667	TGTCCTCGAGTCTTTGACCAATTTGACTTGTGCTGAAGCTCTTCCGAAAACCCACTGA	1726
Db	1802	TGTCCTCGAGTCTTTGACCAAGCTTGAACAGCTTGAACGGCTCCAGAGCCCCCACTGC	1851
Qy	1727	GGGGCTGTGTCTGTGTGACCACTCAAGGGGCTTGAAGAGCCCCCTCGAAGGGCCCTG	1780
Db	1862	GGGGCTGTGTCTGTGTGACCACTCAAGGGGCTTGAAGAGCCCCCTCGAAGGGCCCTG	1862
Qy	1781	-----GTCTGAACGATGTGTGAGGAGCAATCTTGGGGGATAGCTGTCTTCTGTACAG	1834
Db	1922	GAGCTGCTTGAAGGCGCTGGGGGAGAGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG	1981
Qy	1835	ACTGCCAAGAGGTGACTGACGCTTACAGACCAAGCCTTAGAATCTGCTCAAGCTCAGC	1893
Db	1982	ACTGCCAAGAGGTGACAGCGACTTACGACAGGCACTGAGGGTCTGCTCAAGCTCACC	2040

QY 1652 GTGAGACAGATCTGTCTCCGAGTCTTTGACCAATTGACCTTGCTGAACGCTTC 1711
Db 1709 GCGAGACAGATCTGTCTCTGAGTCTTTGACAGAGTGAATTGCTGAACGCTTC 1768
QY 1712 CCGAAACCCCACTGAGGGGCTGTGTCTGTGACCAACTGAGGGGCTTGACAGCTTC 1771
Db 1769 CAGAGCCCACTGAGGGGCTGTGTCTGTGACCAACTGAGGGGCTTGACAGCTTC 1828
QY 1772 CCTCAGAG-----GTCTGAGAGGATGATGAGGAGAAATCCTTGAGGGATAGCT 1819
Db 1829 CCTCAGAGGGGCTCAGAGAGCTGCTGAGGAGCTGAGGAGCTGAGGAGATCTTGAGGGAGCT 1888
QY 1820 GCTTTTCTGACAGATGCTCCAGAGGATGACCTGAGCTCAGACCAAGCCCTAGAAATCT 1879
Db 1889 GCTTTTCTGACAGATGCTCCAGAGGATGACCTGAGCTCAGACCAAGGAGCTCT 1948
QY 1880 GCTCAAGCTCAGC 1893
Db 1949 GCTCAAGCTCAGC 1962

RESULT 11

US-10-322-281-25
; Sequence 25, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945200100
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-25

Query Match 71.7%; Score 1357; DB 18; Length 3404;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

QY 47 CTTCCGCTGCTGCTGAGAGCGCCCGGCTGAGGAGAGGGCTCATCAAGTGGCTA 106
Db 162 CCACTCTCTGCGCGAGAGCTAGCTCGGCTGAGGAGGGCTGATCAAGTGGCTCA 221
QY 107 AACCTGATGAGAGAGAGCGGCTGAGAGCGGACATCAAAACACAACTGCGGAC 166
Db 222 AGCCCTGATGAGAGAGAGCGGCTGAGAGCGGACATCAAAACACAACTGCGGAC 281
QY 167 GCTACGAGTCTGAGAGCGCTGAGAGCGGACCTACGGAGAGTGAAGAGGACAG 226
Db 282 GCTACGAGTCTGAGAGCGCTGAGAGCGGACCTACGGAGAGTGAAGAGGCGG 341
QY 227 AGAGCTGAGGAGCTGTGTGAGCTCAAGTCTCAAGAGAGAGAGAGAGAGAG 286
Db 342 AGAGCTGAGGAGCTGTGTGAGCTCAAGTCTCAAGAGAGAGAGAGAGAGAG 401
QY 287 AGAGCTGAGGAGCTGTGTGAGCTCAAGTCTCAAGAGAGAGAGAGAGAGAG 346
Db 402 AGAGCTGAGGAGCTGTGTGAGCTCAAGTCTCAAGAGAGAGAGAGAGAGAG 461
QY 347 TCATTCATTCATGAGAGTGTGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 406
Db 462 TCATTCATTCATGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 521
QY 407 CAGGCGAGGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 466
Db 522 CAGGCGAGGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 581

QY 467 CAGGCAATTTCTTCCAGACAGATGCTGTGCTCTGACACTACTGCCACAGAGGAGATG 526
Db 582 CTAGGCAATTTCTTCCAGACAGATGCTGTGCTCTGACACTACTAGGCAATGAGAGAGATG 641
QY 527 TTGACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586
Db 642 TTGACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 701
QY 587 CTGACTTTGAGCTCTCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 646
Db 702 CTGACTTTGAGCTCTCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 761
QY 647 GCTCTCTTACGCTGCTGAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 706
Db 762 GCTCTCTTACGCTGCTGAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 821
QY 707 ACAGCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
Db 822 ACAGCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
QY 767 GCGAGATCATTAAG 826
Db 882 GCGAGATCATTAAG 941
QY 827 AGCCGCTGAG 886
Db 942 AACCTCTGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1001
QY 887 CCAACATGAG 946
Db 1002 CCAACATGAG 1061
QY 947 GCGAG 1006
Db 1062 GAG 1121
QY 1007 TGCGGAG 1066
Db 1122 TGCGGAG 1181
QY 1067 GCTTCTTCAAG 1126
Db 1182 GCTTCTTCAAG 1241
QY 1127 ATTCTCTTAAG 1186
Db 1242 ATTCTCTTAAG 1301
QY 1187 CTGAGAGATACCTCTTCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
Db 1302 CTGAGAGATACCTCTTCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 1247 AGAAG 1306
Db 1362 AGAAG 1421
QY 1307 TGCGTGTATCTCAAG 1366
Db 1422 TGCGTGTATCTCAAG 1481
QY 1367 TTAAG 1426
Db 1473 TTAAG 1532
QY 1427 GCGAG 1486
Db 1533 GCGAG 1592
QY 1487 CAG 1546
Db 1593 CAG 1652
QY 1547 CCGGCAAG 1606

Db 1653 CCAGAGAGCTTGGAGCTCGCGGCCCCCACTTCCGCTCCCTGGATGAATCCGCC 1712
Qy 1607 CTTCCATCTGACGCGCGCCCAAGCCCTTCAAGGCTGTGAGAGACAGATCC 1666
Db 1713 CACCTCCGCCCTTGCGCGGCGCAGCCGACCTCAGGGGCTGTGAGGAGAGACATCC 1772
Qy 1667 TGTCTCCGAGTCTTTTGAACCAATTGACCTTGGCAACGTCTTCCGAAACCCCATCTGA 1726
Db 1773 TGTCTTGAAGTCTTTTGAACCAAGCTTGGCAACGTCTTCCGAAACCCCATCTGC 1832
Qy 1727 GGGGCTGTGTGTGTGAGACCACTGAGGGGGCTTGAAGAGCTCCCTCAGAG----- 1780
Db 1833 GGGGCTGTGTGTGTGAGACCACTGAGGGGGCTTGAAGAGAGCCCTCAGAGGGCCCTG 1892
Qy 1781 -----GTGAAAGCATGTGTGAGAGAAATCTTGGGGGATAGCTCTTTCTGTGACG 1834
Db 1893 GAAAGCTCTGAGGCGCTGCGGGGAGGATCTTGGGGGACAGCTGCTTTCCCTGACAG 1952
Qy 1835 ACTGCCAAGAGATGATCTGACAGCTTACAGACCAAGCCCTTGAAGATCTGTCAAGCTCAGC 1893
Db 1953 ACTGCCAAGAGATGATCTGACAGCTTACAGACCAAGGATCTGTCTCAAGCTCAGC 2011

RESULT 12

US-10-618-941-11
; Sequence 11, Application US/10618941
; Publication No. US2004019792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618, 941
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395, 632
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 3463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-11

Query Match 71.7%; Score 1357; DB 18; Length 3463;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

Qy 47 CTTCCGCTGCTGCTCGAGAGCGCCCGCGCTGCGAGCGGCTCATCAAGTCGCTTA 106
Db 223 CCACTCTCTGCGCGCAGACTAGCCCGCGCTGCGAGAGGCTGATCAAGTCGCCCA 282
Qy 107 AACCTGATGAAGAGACAGCGGTGAAGGCGCACTCAAAACAACCTGCGGAC 166
Db 283 AGCCCTTAATGAAGAGACAGCGGTGAAGGCGCACTCAAAACAACCTGCGGAC 342
Qy 167 GCTACGAGTCTGAGAGCGCTGCGCAAGGCACTTACGGGAAGGTGAAGAGGCAAG 226
Db 343 GCTACGAGTCTGAGAGCGCTGCGCAAGGCACTTACGGGAAGGTGAAGAGGCGG 402
Qy 227 AGAGCTCGGGGCGCTGTGTGAGCATCAAGTCATCAGAGAAACAAATCAAGATAGC 286
Db 403 AGAGCTCGGGGCGCTGTGTGAGCATCAAGTCATCAGAGAAACAAATCAAGATAGC 462
Qy 287 AGATCTGCTGCACTATCGAGAGGAGATGAGATCAATGTCTTCACTCAACCCCA 346
Db 463 AAGATCTGATGCACTATCGAGAGGAGATGAGATCAATGTCTTCACTCAACCCCA 522
Qy 347 TCATTCGATCTGATGAGATGTTTGAAGATGAGCAAGATGTTGATGATGAGATG 406
Db 523 TCATTCGATCTGATGAGATGTTTGAAGATGAGCAAGATGTTGATGATGAGATG 582

Qy 407 CCAAGCGAGGATCTGTATGATTAATCATGATGAGCGGCCAGCGCTGATGAGCGGACG 466
Db 583 CCAAGCGGCGGACCTTTATGACTACATGACGAGCGCGCAGAGCTGATGAGCGGACG 642
Qy 467 CCAAGGATTTCTTCCGACATAGGTGTCTCCGCACTTCTGCACTGCGCACAGAGGATG 526
Db 643 CTAGGATTTCTTCCGACATAGGTGTCTCCGCACTTCTGCACTGCGCACAGAGGATG 702
Qy 527 TTCAACGAGATCTCAAGCTGAAAACATCTTCTAGATGCAATGAAATCAAGATG 586
Db 703 TTCAACGAGATCTCAAGCTGAAAACATCTTCTAGATGCAATGAAATCAAGATG 762
Qy 587 CTGACTTGGCTCTTCAACCTGTGACCAAGGCAAGTTCTTCAAGCTTCTGTGAG 646
Db 763 CTGACTTGGCTCTTCAACCTGTGACCAAGGCAAGTTCTTCAAGCTTCTGTGAG 822
Qy 647 GCCCTCTTACGCGCTGCGCAAGATGATCAAGGAAAGCCCTTATGTGGGCGAGAGTGG 706
Db 823 GCCCTCTTATGCGCTGCGCAAGATGATGAAAGCCCTTACACAGGCCAGAGGTGG 882
Qy 707 ACAAGTGTCTTGTGGCGTCTCTGTACATCTGTGTGATGAGCAACATGCTTGTAGC 766
Db 883 ACAAGTGTCTTGTGGCGTCTCTGTACATCTGTGTGATGAGCAACATGCTTGTATG 942
Qy 767 GGCAGATCTATAAACAATGATGAGCAATCACTAACGCGGCTTACCTGAGCGGCCCA 826
Db 943 GGCATGACATAAAGATCTTGTGAAACATGATGAGCAACGCGGCTTACCGGAGCACTTA 1002
Qy 827 AGCGTCCGATGCGCTGCGCTGATCCGGGCGCTTAAATGTGAAACCCCAACCGTGG 886
Db 1003 AACCTGTATGCGCTGCGCTGATCCGGGCGCTTAAATGTGAAACCCCAACCGTGG 1062
Qy 887 CCAACTGAGAGATGATGACAGTCAATGTGTGAGTCACTGAGGTTACACCAAGGATG 946
Db 1063 CCAACTGAGAGATGATGACAGTCACTGAGGTTACACCAAGGATGAGGATG 1122
Qy 947 GGAAGACAGAAAGCTGCGTGAAGGTGAGCACTTATGTGATCTTGGCGGCGCTTCA 1006
Db 1123 GAGAGCAGAGAGCTCGCATGAGAGGTGAGCACTTATGTGATCTTGGCGGCGCTTCA 1182
Qy 1007 TGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
Db 1183 TGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1242
Qy 1067 GCTTCTTCAAGACAGAGTCCGAGAGGTGAGAACTGATCTGAGCTGAGCGGAC 1126
Db 1243 GCTTCTTCAAGACAGAGTCCGAGAGGTGAGAACTGATCTGAGCTGAGCGGAC 1302
Qy 1127 ATTCTCTTAAGAGTCCGAAAGAGATGATGATGATGATGATGATGATGATGATG 1186
Db 1303 ATTCTCTTAAGAGTCCGAAAGAGATGATGATGATGATGATGATGATGATGATG 1362
Qy 1187 CTGAGGATACCTCTTCTGCGCGCTGCGAGAGCGCTTAAAGCTTCCGAAAGGATTTCA 1246
Db 1363 CTGAGGATACCTCTTCTGCGCGCTGCGAGAGCGCTTAAAGCTTCCGAAAGGATTTCA 1422
Qy 1247 AGAAAAAGTCTTACTCTGCTCAGGGAGGTATACAGAGAGCTTACAGAACTCAGACCG 1306
Db 1423 AGAAAAAGTCTTACTCTGCTCAGGGAGGTATACAGAGAGCTTACAGAACTCAGACCG 1482
Qy 1307 TGCCTGATCTTCCAGGAGCGCTGTCTCTGCTGATCTCTGCTTCCAGAGAAAGCATCC 1366
Db 1483 TGCCTGATCTTCCAGGAGCGCTGTCTCTGCTGATCTCTGCTTCCAGAGAAAGCATCC 1533
Qy 1367 TTAAGAAAGTCTGACAGCGGTGAATCTGTGTTACTTCTTCCAGAGCGGAGAGCTG 1426
Db 1534 TCAAGAAAGTCTGACAGCGGTGAATCTGTGTTACTTCTTCCAGAGCGGAGAGCTG 1593
Qy 1427 GGAAGCTCTTGAAGCGCAAGTGTGTTGTGAGTGGAGAGCCCGTGGAGCAGAGCTC 1486
Db 1594 GGAAGCTCTTGAAGCGCAAGTGTGTTGTGAGTGGAGATCCCAAGAGCAGAGCTC 1653

QY 1487 CATAGGTTGAGGCTCTCTCTCCAGCCGCAAGGCAATTCCTCAATGAGCAAGTTCT 1546
DB 1654 CGAAGCTTCAAGGCTGCTCTCTCATGCAAAAGGCAATCTCAAACTGAGCAAGTTCT 1713
QY 1547 CCGCAGAGCTTGAAGAGCACTACCCCTAGCACTTTGGCTCCCTGAGCACTGAGCT 1606
DB 1714 CCGAGACAGCTTGGAGCTGCGGCCCCACACACTTTCGGCTCCCTGAGCACTGAGCT 1773
QY 1607 CTTCCCATCTGACGCGCGCCGCGCCCTCTGAGGAGCTGTGATGAGGACAGATCC 1666
DB 1774 CACTGCGCCCTGCGCGCGCGCGCGCGAGCCCTCAGGGGCTGTGAGGAGGACAGATCC 1833
QY 1667 TGTCTCCAGATCTCTTGAACCAATTTGACTTGAACCTCTTCCGAAACCCCACTGA 1726
DB 1834 TGTCTCTGAGTCTTCTTGAACCAAGTGAACCTTGAACCGGCTCCAGAGCCCACTGC 1893
QY 1727 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGCAAGCTCTCTCAAGAG----- 1780
DB 1894 GGGGCTGTGTGTGTGTGAGCAACCTCAAGGGCTTGAAGAGCCCTCTCAAGAGGGCCTG 1953
QY 1781 -----GTCTGAAGCAATGTGAGGAAATCTTGGGGGATGAGTGTCTTCTGACAG 1834
DB 1954 GAAGCTCTGAGGCGCTGCGCGCGAGATCTTGGGGGAGAGCTGCTTCTCTGACAG 2013
QY 1835 ACTGCCAAGAGTGAAGCTGACAGCTTGAAGCAAGCCCTAGAAATCTCTCAAGCTCAGC 1893
DB 2014 ACTGCCAAGAGTGAAGCTGACAGCTTGAAGCAAGCCCTAGAAATCTCTCAAGCTCAGC 2072

RESULT 13

US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dmanac, Radolje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids at
; TITLE OF INVENTION: PolyPeptides
; FILE REFERENCE: 803 1CNDP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_Fl_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 71.6%; Score 1355.4; DB 17; Length 2501;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CTTCCGCTGAGGCTCTGAGAGCGCCGCGCTGAGGAGCGGGCTATCAATGCGCTTA 106
DB 140 CCACTCCCTCGCGCGAGAGGTACCCGCGCTGAGGAGCGGGCTATCAATGCGCTTA 199
QY 107 AACCTGATGAGAGCAAGCGGCTGAGAGCGGCACTATCAAAACACAACTGAGGACC 166
DB 200 AGCCCTTAATGAGAGAGCGGCTGAGAGCGGCACTATCAAAACACAACTGAGGACC 259
QY 167 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTTACGGAGAGTGAAGAGGACGAG 226

DB 260 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTTACGGAGAGTGAAGAGGCGCGG 319
QY 227 AGAGCTGGGGCGCTGTGTGCGCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 286
DB 320 AGAGCTGGGGCGCTGTGTGCGCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 379
QY 287 AGAGTCTGTGACATACGAGGAGATTTGAGATCATGTCTTCACTCAACCACTCCACA 346
DB 380 AAGATCTGATGCAATACGAGGAGATTTGAGATCATGTCTTCACTCAACCACTCCACA 439
QY 347 TCAATGCAATCAAGATGTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 406
DB 440 TCAATGCAATCAAGATGTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 499
QY 407 CCAAGCCAGAGCGATCTGTATGATTAATCAATGATGAGCGGCAAGCTGATGAGCGGAGC 466
DB 500 CCAAGCCAGAGCGATCTGTATGATTAATCAATGATGAGCGGCAAGCTGATGAGCGGAGC 559
QY 467 CCAAGCAATTTCTTCCAGAGATGCTGTGCGCTGCACTACTGCAACCAAGAGATG 526
DB 560 CTAGGCAATTTCTTCCAGAGATGCTGTGCGCTGCACTACTGCAACCAAGAGATG 619
QY 527 TTCAACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCGCAATGAGAAATCAAGATG 586
DB 620 TTCAACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCGCAATGAGAAATCAAGATG 679
QY 587 CTGACTTTGAGCTCTTCCAACTGTATCAACAAAGGCAAGTTCTTCCAGAGTTCTGTGGA 646
DB 680 CTGACTTTGAGCTCTTCCAACTGTATCAACAAAGGCAAGTTCTTCCAGAGTTCTGTGGA 739
QY 647 GCCCTCTTCAAGCTCTGAGATGATCAACGAGAAAGCCCTATGAGGCGCCAGAGTGG 706
DB 740 GCCCTCTTCAAGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 799
QY 707 ACAAGCTGTCTCTGAGGAGTTCTCTGTATCAATCTGTGTGATGATGATGATGATG 766
DB 800 ACAAGCTGTCTCTGAGGAGTTCTCTGTATCAATCTGTGTGATGATGATGATGATG 859
QY 767 GGGAGATCAATTAACACTGATGAGAAATCAAGTAAGGAGGCTTACCGTAGGCGGCCA 826
DB 860 GGGAGATCAATTAACACTGATGAGAAATCAAGTAAGGAGGCTTACCGTAGGCGGCCA 919
QY 827 AGCGTCCGATGCGCTGTGAGCTGATCCGATGCTGTTAATGATGAACCCGACCGTGGG 886
DB 920 AACCTCTGATGCTGTGAGCTGATCCGATGCTGTTAATGATGAACCCGACCGGCGG 979
QY 887 CCACACTGAGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATG 946
DB 980 CCACACTGAGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATG 1039
QY 947 GGGAGAGAGAGCGCTGAGAGGAGTGGAGCAACCTTAGTGAATTTGGCGGAGCTTCA 1006
DB 1040 GAGAGAGAGAGCGCTGAGAGGAGTGGAGCAACCTTAGTGAATTTGGCGGAGCTTCA 1099
QY 1007 TGGGAGACTGTTAAGTGTGCTCTCGCGCCCTCTCTGAGAAATGAGCAAGTGTGCA 1066
DB 1100 TGGCTGATGCTGCTCGCGCGCTTCTCTCGCGCCCTCTCTGAGAAATGAGGAGTGTGCA 1159
QY 1067 GCTTCTTCAAGAGCAAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1126
DB 1220 ATTGCTCAAGAGTCCCGAGAGATGAATGAATGAATGAATGAATGAATGAATGAAT 1279
QY 1187 CTGAGGATTAATCTTCTGCGCTTGAAGAGAGCTTGAAGTTCGAAAGGCAATTTCTCA 1246
DB 1280 CTGATGATCACTGCGCATGCTGCTGAGAGAGCAATCTCAAGCTCCCAAGGCAATTTCTCA 1339
QY 1247 AGAAAGATCTCTTCTGCTGAGAGGATGAAGAGATCAAGAGACCTTCAAGAACTCAAGCCG 1306
DB 1340 AGAAAGATCTCTTCTGCTGAGAGGATGAAGAGATCAAGAGACCTTCCGAGCTCAAGCCCA 1399

Db 1242 ATTGCTCAAGAAAGTCCCGCAAGAGAAATGACATGCCCCAGTCTCTCCACAGTACACGG 1301
Qy 1187 CTGAGATACCTCTTCTCGCCCTGGCAAGACAGCTTAAAGCTTCGAAAGGCAATTCTCA 1246
Db 1302 CTGATGACACCTGCCCATCGCCCTGGCAAGACAGCTTAAAGGCAATTCTCA 1361
Qy 1247 AGAAAAGTCTCTTACCTCTGTCAGGGAGGTACAGAGAGACCTTCAGAACTCAGACCGG 1306
Db 1362 AGAAGAAAGTCTGACCTCTGTCAGAGGGGTACAGAGAGACCTTCAGAGCTCAGCCCA 1421
Qy 1307 TGCTGATACCTCCAGGGAGAGCTGTCTCTGCTGATCCCTCTCTCCAGAAAGGCAATCC 1366
Db 1422 TCCCTGGAGCCCGAGGGAGGCTGCC-----CGCTCTCTCCAGAGGCAATTC 1472
Qy 1367 TTAAGAGTCTGACAGAGCTGAATCTGATTAATCTCTCTCCAGAGCCGAGAGTCTG 1426
Db 1473 TCAAGAGCCCGCAGAGGAGAGTCTGCTACTACTCTCTCCAGAGCCGAGAGTCTG 1532
Qy 1427 GGGAACTTTAGACGCGCATGATGTGTGTGTGAGTGGGAGCCCGTGGAGCAAGTCTC 1486
Db 1533 GGGAGCTTTGAGACGAGGCGCAGCGTGTGTGTGAGTGGGATCCCAAGAGCAGAGCTTC 1592
Qy 1487 CACAGGCTTCAGAGGCTCTCTCTCCACGCAAGGAGATCTCAAACTCAATGCAAGTCT 1546
Db 1593 CGAAGCTTCAGAGGCTCTCTCTCCACGCAAGGAGATCTCAAACTCAATGCAAGTCT 1652
Qy 1547 CCGGACAGCTTGAAGGCACTACCCCTAGACCTTGGCTCCCTGGAGCAAGTGGCT 1606
Db 1653 CCGAGACGCTTGAAGTCTCGGGGCCCGCCACACCTTGGCTCCCTGGAGTGAATCTCGCC 1712
Qy 1607 CTTCCCATCTGACGCGCGCCGCGCCGCTCCAGGGCTGTGTAGTGAAGACAGCATCC 1666
Db 1713 CACTCGCGCCCTGAGCGCGCGCGCCGCGCCGCTCCAGGGCTGTGTAGTGAAGACAGCATCC 1772
Qy 1667 TGCTCTCCAGATCTTTGACCAATTTGACCTTGCTGAAGCTTCTCCGAAACCCCACTGA 1726
Db 1773 TGCTCTGAGATCTTTGACCAAGCTTGCTGAAGCTTCTCCGAAACCCCACTGC 1832
Qy 1727 GGGGCTGTGTCTGTGTGACCACTGAGGGGGCTTGAACAGCTCCCTCCAGAG----- 1780
Db 1833 GGGGCTGTGTCTGTGTGACCACTGAGGGGGCTTGAAGAGCCCTCCAGAGGGCCCTG 1892
Qy 1781 -----GTCTGAACGATGTGTGACAGAACTTGGGGATGATGCTTTCTCTGACAG 1834
Db 1893 GAACTGCTGATGAGGCGCTGGGCGCAGATCTTTGGGGGACAGCTGCTTTCCCTGACAG 1952
Qy 1835 ACTGCCAAGAGGTGATGACGCTTACAGACAAAGCTTGAAGATCTGCTCAAGCTCAGC 1893
Db 1953 ACTGCCAAGAGGTGATGACGCTTACAGACAAAGCTTGAAGATCTGCTCAAGCTCAGC 2011

RESULT 15
US-10-343-514-13
; Sequence 13, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBEYRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DEUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13
; LENGTH: 1186
; TYPE: DNA
; ORGANISM: RAT
US-10-343-514-13
Query Match 56.6%; Score 1071.4; DB 18; Length 1186;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1114; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 70 GCGCGCGCGCTGGGGAGCGGCTCATCAAGTCGCTTAACTGTGATGAAGAGCGCG 129
Db 1 GCCGGCGCGCTGGGGAGCGGCTCATCAAGTCGCTTAACTGTGATGAAGAGCGCG 60
Qy 130 GTGAAGCGGCACATCAAAACAACCTGCGGACCGCTACGATTCCTGAGAGCGCTG 189
Db 61 GTGAAGCGGCACATCAAAACAACCTGAGGACCGCTACGATTCCTGAGAGCGCTG 120
Qy 190 GGCAGGCGACCTACGAGAGGTGAAGAGGACGAGAGAGCTTGGGCGCTTGGTGCC 249
Db 121 GGCAGGCGACCTACGAGAGGTGAAGAGGACGAGAGAGCTTGGGCGCTTGGTGCC 180
Qy 250 ATCAAGTCCATCAGAAAGAAACAATAATGAGAGATGAGATGCTGCTGCAATACGAGG 309
Db 181 ATCAAGTCCATCAGAAAGAAACAATAATGAGAGATGAGATGCTGCTGCAATACGAGG 240
Qy 310 GAGATGAGATCATGTCTTCACTCAACAACCCCATCATCTGCTCATGAAAGTCTT 369
Db 241 GAGATGAGATCATGTCTTCACTCAACAACCCCATCATCTGCTCATGAAAGTCTT 300
Qy 370 GAGATGAGCAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 429
Db 301 GAGATGAGCAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy 430 TACATCAGTGAAGCGGCGCACGCTGATGATGATGATGATGATGATGATGATGATGAT 489
Db 361 TACATCAGTGAAGCGGCGCACGCTGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 490 GTGTCTGCGCTGCACTACTGTGCAACAGAGGGAGTCTTCAACGAGATCTCAAGCTGGA 549
Db 421 GTGTCTGCGCTGCACTACTGTGCAACAGAGGGAGTCTTCAACGAGATCTCAAGCTGGA 480
Qy 550 AACATCTCTTGAATGCGCAATGAGAAACATCAAGATGCTGATCTTGGCTCTCAACCTG 609
Db 481 AACATCTCTTGAATGCGCAATGAGAAACATCAAGATGCTGATCTTGGCTCTCAACCTG 540
Qy 610 TACACAAAGGCAAGTCTTCAACAGCTTGTGTGAGAGCCCTCTTACGCTCGCTGAG 669
Db 541 TATCAACAAAGGCAAGTCTTCAACAGCTTGTGTGAGAGCCCTCTTACGCTCGCTGAG 600
Qy 670 ATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGCTCTGGGCGTTC 729
Db 601 ATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGCTCTGGGCGTTC 660
Qy 730 CTGTACATCTGTGTGATGAGCAATGCGCTTGTGAAGGCGAGATCATAAACAATCTGAG 789
Db 661 CTGTACATCTGTGTGATGAGCAATGCGCTTGTGAAGGCGAGATCATAAACAATCTGAG 720
Qy 790 AAGCAATCAATACGAGGCGCTTACCGTGAAGCGCGCAAGCGCTGATGCTGTGGCTG 849
Db 721 AAGCAATCAATACGAGGCGCTTACCGTGAAGCGCGCAAGCGCTGATGCTGTGGCTG 780
Qy 850 ATCCGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
Db 781 ATCCGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 910 CATTGATGATCAATGAGGCTTGAACACACCGGATGCGGGAACAGAAAGCCCTGCTGAG 969
Db 841 CATTGATGATCAATGAGGCTTGAACACACCGGATGCGGGAACAGAAAGCTCTGAGAG 900
Qy 970 GGTGGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1029
Db 901 GGTGGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960

QY 1030 TCGGCGCCCTCTGGAATGAGCCAGGTGTGCACTTCTCAAGCAGCAGTGCCG 1089
Db 961 TCCGCCCCCTCTGAGAAATGAGCCAAAGTGTGTAGCTTCTTCAGCAGCATGTGCCG 1020
QY 1090 GAGGTGGAAGCACTGTACTGTGGCTGAGCGGCAACATTTCTTAAGAGTCCGAAAG 1149
Db 1021 GAGGTGGAAGCAGGGGACCGGGGCTGAGCGGCAACATTTCTTAAGAGTCCGCAAG 1080
QY 1150 GAGATGACATGCTCAAAATCTGCAGGTGACCCGGCTGAGGATACCTTTTCGCCCT 1209
Db 1081 GAGATGACATGCTCAAGCTCTGAGATGACCCAGTTGAAGATACCTTCTCGCCCT 1140
QY 1210 GCGAAGAGCAGCCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAAG 1254
Db 1141 GCGAAGAGCAGCCTCAAGCTTCCGAAAGGTATCTCAAGAAAAAG 1185

Search completed: May 16, 2005, 09:29:19
Job time : 1084.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 11, 2005, 14:12:18 ; Search time 64.9672 Seconds
(without alignments)

1994.176 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGTGKVKAKRE.....VNPTRRATLEDVASHMVMVM 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342	100.0	631	2 Q9DBV0	Q9DBV0 m mus muscu
2	1342	100.0	631	2 Q8CIC0	Q8CIC0 mus musculu
3	1338	99.0	639	2 Q80ZM3	Q80ZM3 mus musculu
4	1328	99.0	639	2 Q8BZM4	Q8BZM4 mus musculu
5	1313	97.8	630	2 Q66HE5	Q66HE5 rattus norv
6	1293	96.3	628	2 Q9H093	Q9H093 homo sapien
7	1128.5	84.1	661	1 ARK5 HUMAN	ARK5 HUMAN
8	1120.5	83.5	658	2 Q64IK5	Q64IK5 mus musculu
9	1025	76.4	575	2 Q616D6	Q616D6 mus musculu
10	824.5	61.4	383	2 Q7QBH4	Q7QBH4 anopheles g
11	822.5	61.3	1180	2 Q7KSS0	Q7KSS0 drosophila
12	822.5	61.3	1180	2 Q7YU19	Q7YU19 drosophila
13	822.5	61.3	1427	2 Q9VH05	Q9VH05 drosophila
14	813.5	60.6	1551	2 Q65ZH3	Q65ZH3 caenorhabdi
15	813.5	60.6	1592	2 Q65ZH3	Q65ZH3 caenorhabdi
16	665.5	49.6	752	1 MKK4 HUMAN	MKK4 HUMAN
17	665.5	49.6	752	2 Q8CIP4	Q8CIP4 mus musculu
18	662.5	49.2	752	2 Q8NG37	Q8NG37 homo sapien
19	660.5	48.6	725	2 Q804T2	Q804T2 xenopus lae
20	651.5	48.5	729	2 Q9JKE4	Q9JKE4 mus musculu
21	651.5	48.5	729	2 Q9JKE5	Q9JKE5 mus musculu
22	650	48.4	712	2 Q7QBH3	Q7QBH3 anopheles g
23	647.5	48.2	691	2 Q96RG0	Q96RG0 homo sapien
24	647.5	48.2	719	2 Q96RG0	Q96RG0 homo sapien
25	647.5	48.2	722	2 Q08679	Q08679 rattus norv
26	647.5	48.2	722	2 Q6PR4	Q6PR4 mus musculu
27	647.5	48.2	745	2 Q15524	Q15524 homo sapien
28	647.5	48.2	755	2 Q7KZ17	Q7KZ17 homo sapien
29	647.5	48.2	778	2 Q96HB3	Q96HB3 homo sapien
30	647.5	48.2	797	2 Q8VHFO	Q8VHFO rattus norv
31	647.5	48.2	888	2 Q8BR95	Q8BR95 mus musculu

32	644	48.0	512	2 P92958	P92958 arabidopsis
33	644	48.0	512	2 P92968	P92968 arabidopsis
34	643.5	48.0	722	2 Q802W0	Q802W0 brachydanio
35	642.5	47.9	792	2 Q61WT7	Q61WT7 xenopus lae
36	642.5	47.9	793	2 Q08678	Q08678 rattus norv
37	642.5	47.9	795	2 Q8VHJ5	Q8VHJ5 mus musculu
38	641.5	47.8	776	2 Q7ZYL7	Q7ZYL7 xenopus lae
39	641.5	47.8	785	2 Q8QGV3	Q8QGV3 xenopus lae
40	641	47.8	504	2 P93113	P93113 cucumis sat
41	640.5	47.7	462	2 Q75271	Q75271 homo sapien
42	640.5	47.7	542	2 Q6V8Y5	Q6V8Y5 physcomitre
43	639.5	47.7	780	2 Q804T1	Q804T1 xenopus lae
44	638.5	47.6	1398	2 Q7Z268	Q7Z268 drosophila
45	638.5	47.6	1398	2 Q9W532	Q9W532 drosophila

ALIGNMENTS

RESULT 1	Q9DBV0	PRELIMINARY;	PRT;	631 AA.
AC	Q9DBV0;			
DT	01-JUN-2001 (TrEMBLrel. 17, last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, last annotation update)			
DT	05-JUL-2004 (TrEMBLrel. 27, last annotation update)			
DE	Mus musculus adult male lung cDNA, RIKEN full-length enriched library,			
DE	clone:1200013B22 product:weakly similar to PROBABLE SERINE/THREONINE-			
DE	PROTEIN KINASE KIAA0537 (EC 2.7.1.-) (Mus musculus adult male cecum			
DE	cDNA, RIKEN full-length enriched library, clone:9130215K18			
DE	product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE			
DE	KIAA0537).			
GN	Name=1200013B22R1k;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;			
RA	Carninci P., Hayashizaki Y.;			
RT	"High-efficiency full-length cDNA cloning.";			
RT	Methods. Enzymol. 303:19-44(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;			
RA	RIKEN FANTOM Consortium;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RT	Nature 409:685-690(2001).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RA	The FANTOM Consortium;			
RT	"The RIKEN Genome Exploration Research Group Phase I & II Team:			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs.";			
RT	Nature 420:563-573(2002).			
RL	[4]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;			
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,			
RA	Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;			
RT	"Normalization and subcloning of cap-trapper-selected cDNAs to			
RT	prepare full-length cDNA libraries for rapid discovery of new genes.";			
RT	Genome Res. 10:1617-1630(2000).			
RL	[5]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;			
RX	MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;			
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,			

RA Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-38-format
 RT sequencing pipeline with 384 multicapillary sequencer.",
 RL Genome Res. 10:11757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cecum;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirokane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ono M., Okazaki Y.,
 RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK004937; BAB23518.1; -
 DR EMBL; AK033672; BAC28421.1; -
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR0011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 631 AA; 69806 MW; D51C042DE6CC174B CRC64;

Query Match 100.0%; Score 1342; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred. No. 2,6e-97;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGVKKARESSGRLVAIKSRDKIKDEODLHIREIEMSLNHPH 60
 DB 57 YEFLETLGKGTGVKKARESSGRLVAIKSRDKIKDEODLHIREIEMSLNHPH 116
 QY 61 IAIHEFENSSKIVIMEYASRDLYDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
 DB 117 IAIHEFENSSKIVIMEYASRDLYDYISERPRLSERDARHFFROIYSALHYCHONGIV 176

QY 121 HRDLKLENTLLDANGNIKIADPGLSNLYHKGKFLQFPCCSPLYASPEIYNGKPYGPEVD 180
 DB 177 HRDLKLENTLLDANGNIKIADPGLSNLYHKGKFLQFPCCSPLYASPEIYNGKPYGPEVD 236
 QY 181 SMSLGVLLYLYVGHGTPFGQDHKTIVKQISNGAYEPPKPSDAGLIRWLWVNPERRA 240
 DB 237 SMSLGVLLYLYVGHGTPFGQDHKTIVKQISNGAYEPPKPSDAGLIRWLWVNPERRA 296
 QY 241 TLEDVASHWVWVW 253
 DB 297 TLEDVASHWVWVW 309
 RESULT 2
 ID 08C1C0 PRELIMINARY; PRT; 631 AA.
 AC 08C1C0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 1200013B22Rik protein.
 GN Name=1200013B22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Smealton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidl T.B., Tohyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalsoda U., Smalls D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC033302; AAH33302.1; -
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR0011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 631 AA; 69745 MW; C6C9042247D9174F CRC64;
Query Match 100.0%; Score 1342; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 2, 6e-97;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YEFLETGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 60
DB 57 YEFLETGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 116
QY 61 IAIHEVENSCKIVMEYASRGDLVYISERPRLSERDARHFFROIYVSAH 120
DB 117 IAIHEVENSCKIVMEYASRGDLVYISERPRLSERDARHFFROIYVSAH 176
QY 121 HRDLKLENIILDDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 180
DB 177 HRDLKLENIILDDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 236
QY 181 SMSGLVLLYILVHGTMPPDGDHKTIVKQISNGAYRBPSPDAGLIRWL 240
DB 237 SMSGLVLLYILVHGTMPPDGDHKTIVKQISNGAYRBPSPDAGLIRWL 296
QY 241 TLEDVASHWVNM 253
DB 297 TLEDVASHWVNM 309
RESULT 3
ID 0802W3 PRELIMINARY; PRT; 639 AA.
AC 0802W3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
GN 1200013B22R1k; Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=2328257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson W.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.B.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (PEB-2003) to the EMBL/GenBank/DBJ databases.
CC -/- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; BC046833; AA046833.1; -.
DR HSP; P31751; IGZK.

DR MGI; 1921387; 1200013B22R1k.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004648; P:Protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR007219; Prot. kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR Pfam; PF00069; Kinase; 1.
DR Prodom; PD000001; Prot. kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 639 AA; 70675 MW; FB9C40228F53872C CRC64;
Query Match 99.0%; Score 1328; DB 2; Length 639;
Best Local Similarity 96.9%; Pred. No. 3, 3e-96;
Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 YEFLETGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 60
DB 57 YEFLETGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 116
QY 61 IAIHE-----VFENSCKIVMEYASRGDLVYISERPRLSERDARHFFROIYVSAH 112
DB 117 IAIHEVGRSLVTVFENSSCKIVMEYASRGDLVYISERPRLSERDARHFFROIYVSAH 176
QY 113 YCHQNGIVHRDLKLENIILDDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 172
DB 177 YCHQNGIVHRDLKLENIILDDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 236
QY 173 PYGPEVDSWGLVLLYILVHGTMPPDGDHKTIVKQISNGAYRBPSPDAGLIRWL 232
DB 237 PYGPEVDSWGLVLLYILVHGTMPPDGDHKTIVKQISNGAYRBPSPDAGLIRWL 296
QY 233 MNVPTRRATLEDVASHWVNM 253
DB 297 MNVPTRRATLEDVASHWVNM 317
RESULT 4
ID 08B2N4 PRELIMINARY; PRT; 639 AA.
AC 08B2N4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:933015AN24 product:weakly similar to PROBABLE
DE SERINE/THREONINE-PROTEIN KINASE KIA00537 (BC 2.7.1.1.-).
GN Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:19-44(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
[3]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA The PANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Aikawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Morimatsu M., Hayashizaki Y.,
RT Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
CC -1- Similarity: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AK034082; BAC26575.1; -.
DR HSSP; P31751; IGZK.
DR MGD; MGI:1921387; 1200013B22Rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002230; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 639 AA; 70632 MW; F89BEBE915C95F85D CRC64;
Query Match 99.0%; Score 1328; DB 2; Length 639;
Best Local Similarity 96.9%; Pred. No. 3.3e-96;
Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRDKIKDEBDLHIREIRIMSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRDKIKDEBDLHIREIRIMSLNHPH 116

QY 61 IAIHE-----VFENSSKIYVMEYASRGDIYDISEPRUSEDPARHFFQIYSALH 112
DB 117 IAIHEVGRSLYVVFENSSKIYVMEYASRGDIYDISEPRUSEDPARHFFQIYSALH 176
QY 113 YCHQNGIVHRDLEKLEMLIDANGNITADPGSNLHKKFKLOTFCGSPLYASPELVNGK 172
DB 177 YCHQNGIVHRDLEKLEMLIDANGNITADPGSNLHKKFKLOTFCGSPLYASPELVNGK 236
QY 173 PYGPEVDSWSLGLVLLIYVHGTPPDGQDHKTLYQISNGAYREPPKSDACGLIRMLL 232
DB 237 PYGPEVDSWSLGLVLLIYVHGTPPDGQDHKTLYQISNGAYREPPKSDACGLIRMLL 296
QY 233 MNVPTRRATLEDVASHWVNW 253
DB 297 MNVPTRRATLEDVASHWVNW 317
RESULT 5
ID 066HE5 PRELIMINARY; PRT; 630 AA.
AC 066HE5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang Y., Hsieh F.,
RA Diatchenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalón D.K., Munz D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC081899; AAH81899.1;
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002230; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.

SQ SEQUENCE 630 AA; 69952 MW; 1626427ABCD5F66E CRC64;
 Query Match 97.8%; Score 1313; DB 2; Length 630;
 Best Local Similarity 98.0%; Pred. No. 5e-95;
 Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKI KEDODLHIREIEMSLNPHI 60
 DB 57 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKI KEDODLHIREIEMSLNPHI 116
 QY 61 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 120
 DB 117 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 176
 QY 121 HRDLKLENTLLDANGNKIKIADFGLSNLYHGKFLQTFGSGPLVASEIYVNGKPYGPEVD 180
 DB 177 HRDLKLENTLLDANGNKIKIADFGLSNLYHGKFLQTFGSGPLVASEIYVNGKPYGPEVD 236
 QY 181 SMSLGVLLYLTVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRMLMVNPTERRA 240
 DB 237 SMSLGVLLYLTVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRMLMVNPTERRA 296
 QY 241 TLEDVASHMWVNW 253
 DB 297 TLEDVASHMWVNW 309

RESULT 6
 ID Q9H093 PRELIMINARY; PRT; 628 AA.
 AC Q9H093;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein DKFZp434J037 (Hypothetical protein SNARK)
 GN Name=DKFZp434J037; Synonyms=SNARK;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RG The German cDNA Consortium;
 RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Fodor G., Han M., Wiemann S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidi T.B., Tomihata S., Carninci P., Prange C.,
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
 RA Krzywinski M.I., Skalek U., Smillie D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Itoigai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hito Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.;
 RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- Similarity: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AL136891; CAB6825.1; -;
 DR EMBL; BC017306; AAH17306.1; -;
 DR EMBL; AK074830; BAC11234.1; -;
 DR HSSP; P1751; 1GZK.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot. kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR008271; Ser_Thr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR Prodom; PD000001; Prot. kinase; 1.
 DR SMART; SM00220; S_TKc_1.
 DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 628 AA; 69611 MW; F76F8B1BF9AF4C7 CRC64;

Query Match 96.3%; Score 1293; DB 2; Length 628;
 Best Local Similarity 95.7%; Pred. No. 1.9e-93;
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKI KEDODLHIREIEMSLNPHI 60
 DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKI KEDODLHIREIEMSLNPHI 112
 QY 61 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 120
 DB 113 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 172
 QY 121 HRDLKLENTLLDANGNKIKIADFGLSNLYHGKFLQTFGSGPLVASEIYVNGKPYGPEVD 180
 DB 173 HRDLKLENTLLDANGNKIKIADFGLSNLYHGKFLQTFGSGPLVASEIYVNGKPYGPEVD 232
 QY 181 SMSLGVLLYLTVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRMLMVNPTERRA 240
 DB 233 SMSLGVLLYLTVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRMLMVNPTERRA 292
 QY 241 TLEDVASHMWVNW 253
 DB 293 TLEDVASHMWVNW 305

RESULT 7
 ARKS HUMAN
 ID ARKS HUMAN STANDARD; PRT; 661 AA.
 AC 060265;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE AMPK-related protein kinase 5 (RC 2.7.1.37).
 GN Name=ARK5; Synonyms=KIAA0537;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;

RN SEQUENCE FROM N.A.
RC TISSUE=Brain.
RA MEDLINE=9820545; PubMed=9628581;
RA Nagae T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [2]
RP CHARACTERIZATION, AND MUTAGENESIS OF SER-600.
RX MEDLINE=22393479; PubMed=12409306; DOI=10.1074/jbc.M206025200;
RA Suzuki A., Kusakai G.-I., Kishimoto A., Lu J., Ogura T., Lavin M.F.,
RA Esumi H.;
RT "Identification of a novel protein kinase mediating Akt survival
RT signaling to the ATM protein.";
RL J. Biol. Chem. 278:48-53(2003).
CC -1- FUNCTION: Involved in tolerance to glucose starvation.
CC Phosphorylates ATM.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: Activated by PKA/AKT1 during glucose
CC starvation.
CC -1- TISSUE SPECIFICITY: Expressed at high levels in heart and brain,
CC and at lower levels in skeletal muscle, kidney, ovary, placenta,
CC lung and liver.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011109; BAA25463.1; -.
DR HSSP; P31751; IGZK.
DR MIM; 608130; -.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser thr_pkin_AS.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR KW ATP-binding; Phosphorylation; Polymorphism;
DR Serine/threonine-protein kinase; Transferase.
FT DOMAIN 55 306 Protein kinase.
FT NP BIND 61 69 ATP (By similarity).
FT ACT_SITE 178 178 Proton acceptor (By similarity).
FT BINDING 84 84 ATP (By similarity).
FT MOD_RES 600 600 Phosphoserine (By PKB/AKT1).
FT VARIANT 543 543 P -> R (in dbSNP:3741883).
FT FTID=VAR_017246.
FT MUTAGEN 600 600 S->A: No phosphorylation.
SQ SEQUENCE 661 AA; 74304 MW; 806F37D52CA4718F CRC64;
Query Match 84.1%; Score 1128.5; DB 1; Length 661;
Best Local Similarity 82.3%; Pred. No. 1.8e-80; Indels 1; Gaps 1;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;
QY 1 YEFLETLGKGTGKVKARBS-SGRVIAIKSRKDKIKOEODLHRRREIEMSLNHPH 59
DB 55 YELQETLGKGTGKVKARERFSGRVAIAIKSRKDKIKOEODVWHRRREIEMSLNHPH 114
QY 60 IAIHVEFENSKIVTVMEYASRGDIYDISSRPRLSERDARHFFQIYSALHYCHQNGI 119
DB 115 IISIVVFENKDKIVIMEYASRGELDYDISSRRRLSERTRHFFQIYSAVHYCHQNGV 174
QY 120 VHRDLKLENIILDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGRPYGPV 179

DB 175 VHRDLKLENIILDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGRPYGPV 234
QY 180 DSWSLGVLIIYLVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRMLMNPTRR 239
DB 235 DSWALGVLLIYLVYGTMPDGFHDKILRIQISSGEYREPTQPSDARGLIRMLMNPDRR 294
QY 240 ATLEDVASHRMVNM 253
DB 295 ATLEDIANHWMVNM 308
RESULT 8
ID 0641K5 PRELIMINARY; PRT; 658 AA.
AC 0641K5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE RIKEN cDNA B230104P22.
GN Name=B230104P22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klanner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smaltz D.E., Schenck A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082328; AAH82328.1; -.
SQ SEQUENCE 658 AA; 73661 MW; E7176F13B75B889F CRC64;
Query Match 83.5%; Score 1120.5; DB 2; Length 658;
Best Local Similarity 81.5%; Pred. No. 7.9e-80;
Matches 207; Conservative 25; Mismatches 21; Indels 1; Gaps 1;
QY 1 YEFLETLGKGTGKVKARBS-SGRVIAIKSRKDKIKOEODLHRRREIEMSLNHPH 59
DB 56 YELQETLGKGTGKVKARERFSGRVAIAIKSRKDKIKELDMVWHRRREIEMSLNHPH 115
QY 60 IAIHVEFENSKIVTVMEYASRGDIYDISSRPRLSERDARHFFQIYSALHYCHQNGI 119
DB 116 IISIVVFENKDKIVIMEYASRGELDYDISSRRRLSERTRHFFQIYSAVHYCHQNGV 175
QY 120 VHRDLKLENIILDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGRPYGPV 179
DB 176 VHRDLKLENIILDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGRPYGPV 235
QY 180 DSWSLGVLIIYLVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRMLMNPTRR 239

Db 236 DSWALGVLLYTLTYGMPDPGFDHKNLIRQISGSEYREPTQPSDAGLIRMLMNVNPDNR 295
QY 240 ATLEDVASHMWNM 253
Db 296 ATLEDIANMWNM 309

RESULT 9
ID 061D6 PRELIMINARY; PRT; 575 AA.
AC 061D6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DE MKIA0537 protein (Fragment).
GN Name=mkIA0537;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Iamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Hasegawa T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIA Gene:
RT III. The Complete Nucleotide Sequences of 500 Mouse KIA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-fractionated Libraries.";
RL DNA Res. 10:167-180(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Nagase T., Ohara O., Koga H.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC EMBL; AB182364; BAD2395.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase, like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 575 AA; 64949 MW; 3CA0796FBA199D2B CRC64;

Query Match 76.4%; Score 1025; DB 2; Length 575;
Best Local Similarity 82.3%; Pred. No. 2.3e-72;
Matches 186; Conservative 23; Mismatches 17; Indels 0; Gaps 0;

QY 28 IKSIRKDKIKDEODLHIREIREIMSLNPHPIAIEHEVENSKIYIVMEYASRGDLYD 87
Db 1 IKSIRKDKIKDEODLHIREIREIMSLNPHPIAIEHEVENSKIYIVMEYASRGDLYD 60
QY 88 YISERPRISERDARHFPROIVSALAYCHONGIVHRDLKLENIILLDANGIKIADFGLSNL 147
Db 61 YISERPRISERDARHFPROIVSALAYCHONGIVHRDLKLENIILLDANGIKIADFGLSNL 120
QY 148 YHKGKFLQTFCCGSLYASPEIVNGKPYGVDSWMSGLVLLYLVHGTMPEFDQDKTLY 207
Db 121 YQDKKFLQTFCCGSLYASPEIVNGKPYGVDSWMSGLVLLYLVHGTMPEFDQDKTLY 180
QY 208 KOISNGAYREPPKPSDAGLIRMLMNVNPDNRATLEDVASHMWNM 253

Db 181 KOISGSEYREPTQPSDAGLIRMLMNVNPDNRATLEDIANMWNM 226
RESULT 10
ID 070BH4 PRELIMINARY; PRT; 383 AA.
AC 070BH4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Eb1p3238 (Fragment).
GN Name=Eb1G3238; ORFNames=ENSGANG00000012297;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RC Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008879; EAA08421.1; -.
DR HSSP; P31751; IGZK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase, like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
FT NON_TER 383
SQ SEQUENCE 383 AA; 43133 MW; 294732D45CFC1023 CRC64;

Query Match 61.4%; Score 824.5; DB 2; Length 383;
Best Local Similarity 61.3%; Pred. No. 9e-57;
Matches 155; Conservative 44; Mismatches 53; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVKKA-RESSGRVIAIKSRDKIKDEODLHIREIREIMSLNPH 59
Db 44 FDIKTLGGGTGKQVGLINKETGCEVALTKIKSKIEFADLRIRRVQIMSSVQHRN 103
QY 60 IIAIHEVENSKIYIVMEYASRGDLYDYISERPRISERDARHFPROIVSALAYCHONGI 119
Db 104 IIAIHEVENSKIYIVMEYASRGDLYDYISERPRISERDARHFPROIVSALAYCHONGI 163
QY 120 VHRDLKLENIILLDANGIKIADFGLSNLVHKGKFLQTFCCGSLYASPEIVNGKPYGVDS 179
Db 164 CHRDLKLENIILLDANGIKIADFGLSNLVHKGKFLQTFCCGSLYASPEIVNGKPYGVDS 223
QY 180 DSWMSGLVLLYLVHGTMPEFDQDKTLYVQISNGAYREPPKPSDAGLIRMLMNVNPDNR 239
Db 224 DSWMSGLVLLYLVHGTMPEFDQDKTLYVQISNGAYREPPKPSDAGLIRMLMNVNPDNR 283
QY 240 ATLEDVASHMWNM 252
Db 284 ANIBQICMWNM 296

RESULT 11
ID 07KS50 PRELIMINARY; PRT; 1180 AA.

ID	Q65ZH4	PRELIMINARY	PRT	1551 AA
AC	Q65ZH4			
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)			
QY	1	YEPFETLKGTYGKVKKA-RESSGRLVAIKSIKDKIDKEDDLHIRREIETMSLNPH	59	
DB	70	FDIIKKIGGGYGVQGLGNETGTGEGVAIKTKKIKKIKAEADLVAKREVGQIMSSVHPN	129	
QY	60	IIAIEVEFENSKTIVYMEVARSQDLDYISERPLSRSDAHFPRQIVSALHYCHQNGI	119	
DB	130	IIHIEVEFENRKKWLIVMEFAAGGELDYILSERKVLTEEARRIIPQVATVAYYCHKIKI	189	
QY	120	VHRLPKLENIILLDANGNIKINDPGLSNLYHKGKPLQTCGSPLYASPEIVNGKPYGGEV	179	
DB	190	CHRLPKLENIILLDERGANKIADFGLSNVDQRLTGTCGSPLYASPEIVETPGQGEV	249	
QY	180	DSWISGLVLYILVHGTMPPDQGDHHTLVKQISNGAYREPPEKPSDGGILRWLWNPTRR	239	
DB	250	DSWISGLVLYILVHGTMPPDQGDHHTLVKQISNGAYREPPEKPSDGGILRWLWNPTRR	309	
QY	240	ATLEDVASHMWNV 252		
DB	310	ASIEQICSHMWNV 322		
RT	RESULT 14			
RT	Q65ZH4			
AC	Q65ZH4			
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)			

Dt	25-OCT-104 (TrEMBLrel. 28, last annotation update)
DE	Hypothetical protein B0496.3.
GN	Name=B0496.3; ORFNames=B0496.3;
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC	Rhabditidae; Peloderinae; Caenorhabditis.
OX	NCBI_TaxID=6239;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RG	MEDLINE=99069613; PubMed=9851916;
RT	WormBase Consortium;
RT	"Genome sequence of the nematode C. elegans: a platform for
RL	investigating biology. The C. elegans Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RA	Murray J., Le T.T.;
RT	"The sequence of C. elegans cosmid B0496."
RL	Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RA	Waterston R.;
RL	Submitted (Feb-2003) to the EMBL/GenBank/DBJ databases.
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RG	WormBase Consortium;
RL	Submitted (Aug-2004) to the EMBL/Genbank/DBJ databases.
DR	EMBL; U58749; AAU05597.1; -.
DR	InterPro: IPR011009; Kinase_1like.
DR	InterPro: IPR000719; Prot_Kinase.
DR	InterPro: IPR002290; Ser_thr_kinase.
DR	InterPro: IPR008271; Ser_thr_Pkin_AS.
DR	InterPro: IPR001245; Tyr_Pkinase.
DR	Pfam; PF00069; Pkinase; 1.
DR	ProDom; PD000001; Prot_Kinase; 1.
DR	SMART; SM00219; Tyrcg; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	Hypothetical protein.
SQ	SEQUENCE 1551 AA; 176395 MW; 1FBE62632A44C85 CRC64;
<hr/>	
Query Match	60.6%; Score 813.5; DB 2; Length 1551;
Best Local Similarity	61.7%; Pred. No. 3.4e-55;
Matches 156; Conservative 33; Mismatches 63; Indels 1; Gaps 1	
Qy	1 YEFLELTGAGTGYKVKARERS-GRVIAIKSRDKDIKDODDLHRRRIEIMSLNHPH 59
Db	75 FELTKLGSGGTGYKGLSLAYDHDFREVAVALKIKAISAKSDLVIRIREIRIMSALNHPN 134
Qy	60 IIAIHEVFENSKEIVIMEYASRGLDYDYSERPLSERDAHRPFQIVSALHYCHONGI 119
Db	135 IIQIYVEFENKOKIIIMWEYSSGGLYDVYVRCSGLPEAEARRIFQITSALVYCHKHRY 194
Qy	120 VNRDLKTLENIILLDANGNIKADFGLSNLTHHGKFLQTFCCSPLYASGEINYGPYPGV 179
Db	195 AHHDLKTLENIILLDONNNNAKIADFGLSNFYADKNILITFCPSGLVASPEIINGTPYKGPEV 254
Qy	180 DSGMSGLVLLIYLIVHGMPDPGDQDHTTLVKOISNGAYREPPRPSPDACGLIRMLLVNPTR 239
Db	255 DCSSGLILLIYLLTVYSSMPPDGDDFNRMVQIKRGAYFEPTPTSYASVLINMLRVNERR 314
Qy	240 ATLEDVASHMWVN 252
Db	315 ATIFDLASHMWLN 327

Q65ZH3
ID Q65ZH3 PRELIMINARY; PRT; 1592 AA.
AC Q65ZH3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3.
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peioderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; Pubmed=9851916;
RG Wormbase Consortium;
RT "genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Le T.T.
RT "The sequence of C. elegans cosmid B0496.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
EMBL; US8749; AAU05598.1; -
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR008271; Ser_Thr_kinase.
DR InterPro: IPR01245; Tyr_kinase.
DR Pfam: PF00069; Kinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; Tyr_Kc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KM Hypothetical protein.
SQ SEQUENCE 1592 AA; 180962 MW; E28D6FF934C43F76 CRC64;

Query Match 60.6%; Score 813.5; DB 2; Length 1592;
Best Local Similarity 61.7%; Pred. No. 3.5e-55;
Matches 156; Conservative 33; Mismatches 63; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGYGKARSS--GRVAIKSIRKDKIKDEODLHIREIEMSLNHPH 59
DB 75 FELTKLGSGTYGVSLAYDHKEDREVAVKLIKSALESKADLVRIIRREIRIMSLNHPN 134
QY 60 IIAIHEVFENSKIIVMEYASRGDLVYISERPRLSERDARHPFQIVSALHYCHONGI 119
DB 135 IIQIYEVFENKDKIIVMEYSSGSELDYVSRGSLPEAEARRIFRQITSALVYCHKHRY 194
QY 120 VHRDLKLENTLLDANGNIKADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEV 179
DB 195 AHRDLKLENTLLDQNNNAKIADFGLSNYFADKMLTTFCCGSPLYASPEIINGTPYKGPYV 254
QY 180 DSNLSGLVLYIVGTMFGPDGDHKTLYKQISNGAYRREPSPKSDACGLIRLLMNVNPTRR 239
DB 255 DCSNLSGLVLYIVGTMFGPDGRDPRNRVQIKRGAIFEPETPTASTMLIRMLRVNPERR 314
QY 240 ATLEDVASHWVNVN 252

DB 315 ATLEDVASHWVNVN 327

Search completed: May 11, 2005, 14:24:01
Job time : 65.9672 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:24:11 ; Search time 52.0882 Seconds
(without alignments)
1622.537 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGYGVKAKARE.....VNPTRRATLEDVASHMWVNW 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications_AA:*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	14	US-10-355-975-11 Sequence 11, Appl
2	1325	98.7	251	16	US-10-343-514-103 Sequence 103, Appl
3	1313	97.8	630	16	US-10-343-514-41 Sequence 41, Appl
4	1296	96.6	251	16	US-10-343-514-50 Sequence 50, Appl
5	1293	96.3	628	9	US-09-963-159-2 Sequence 2, Appl
6	1293	96.3	628	15	US-10-423-543-44 Sequence 44, Appl
7	1293	96.3	672	16	US-10-618-941-77 Sequence 77, Appl
8	1286	95.8	594	15	US-10-311-034-112 Sequence 12, Appl
9	1276	95.1	251	16	US-10-343-514-101 Sequence 101, Appl
10	1273	94.9	640	16	US-10-322-281-23 Sequence 23, Appl
11	1230	91.7	616	16	US-10-322-281-26 Sequence 26, Appl
12	1128.5	84.1	661	9	US-09-780-949-2 Sequence 2, Appl
13	1128.5	84.1	661	9	US-09-780-949-6 Sequence 6, Appl

14	1128.5	84.1	661	14	US-10-354-358-82	Sequence 82, Appl
15	1111.5	82.8	252	16	US-10-343-514-102	Sequence 102, Appl
16	814	60.7	530	9	US-09-836-392-20	Sequence 20, Appl
17	785.5	58.5	246	10	US-09-898-8378-29	Sequence 29, Appl
18	665.5	49.6	508	15	US-10-016-248-24	Sequence 24, Appl
19	665.5	49.6	639	15	US-10-016-248-71	Sequence 71, Appl
20	665.5	49.6	639	15	US-10-016-248-72	Sequence 72, Appl
21	665.5	49.6	688	15	US-10-276-645-8	Sequence 8, Appl
22	665.5	49.6	703	15	US-10-016-248-70	Sequence 70, Appl
23	665.5	49.6	752	9	US-09-835-081-2	Sequence 2, Appl
24	665.5	49.6	752	15	US-10-258-106-16	Sequence 16, Appl
25	665.5	49.6	752	15	US-10-276-645-7	Sequence 7, Appl
26	665.5	49.6	769	15	US-10-363-616-403	Sequence 403, Appl
27	665.5	49.6	825	15	US-10-425-114-54516	Sequence 54516, A
28	661.5	49.3	752	16	US-10-618-941-79	Sequence 79, Appl
29	660.5	49.2	639	15	US-10-016-248-73	Sequence 73, Appl
30	660.5	49.2	660	15	US-10-276-645-6	Sequence 6, Appl
31	660.5	49.2	688	14	US-10-161-565-28	Sequence 28, Appl
32	660.5	49.2	688	14	US-10-161-565-29	Sequence 29, Appl
33	660.5	49.2	724	15	US-10-276-645-5	Sequence 5, Appl
34	651.5	48.5	744	9	US-09-835-081-4	Sequence 4, Appl
35	649.5	48.4	729	14	US-10-142-356-11	Sequence 11, Appl
36	649.5	48.4	729	14	US-10-195-101-33	Sequence 33, Appl
37	649.5	48.4	729	14	US-10-161-565-26	Sequence 26, Appl
38	649.5	48.4	744	9	US-09-919-585-3	Sequence 3, Appl
39	647.5	48.2	691	9	US-09-919-585-6	Sequence 6, Appl
40	647.5	48.2	722	15	US-10-274-194-4	Sequence 4, Appl
41	647.5	48.2	722	16	US-10-760-407-4	Sequence 4, Appl
42	647.5	48.2	724	9	US-09-919-585-9	Sequence 9, Appl
43	647.5	48.2	745	14	US-10-195-101-36	Sequence 36, Appl
44	647.5	48.2	745	14	US-10-161-565-24	Sequence 24, Appl
45	647.5	48.2	745	15	US-10-260-708-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-10-355-975-11
; Sequence 11, Application US/10355975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/10/355,975
; PRIOR FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579,664B
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975-11

Query Match 100.0%; Score 1342; DB 14; Length 631;
Best Local Similarity 100.0%; Pred No. 1.2e-100;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YEFLETLGKGYGVKAKARESSGRLVAIKSRKDKIKEDQDLHRRRIEIMSLNPHI 60
DB 57 YEFLETLGKGYGVKAKARESSGRLVAIKSRKDKIKEDQDLHRRRIEIMSLNPHI 116
QY 61 IAIHEVFENSSKIVIMEYASRDLDYIISERRRLSERDARRHFRQIVASLHYCHONGIV 120
DB 117 IAIHEVFENSSKIVIMEYASRDLDYIISERRRLSERDARRHFRQIVASLHYCHONGIV 176

Qy	121	HRDLKLENIILDLANGNIKADFGLSLHYHKGKFLQFCCGSLVLAASPEIYVNGKPYGVEVD	180
Db	177	HRDLKLENIILDLANGNIKADFGLSLHYHKGKFLQFCCGSLVLAASPEIYVNGKPYGVEVD	236
Qy	181	SMSLGVLIIYLIVHGMPFPGDGHKTLVKQISNGAYRBPKPSPACGLIRLMLVNPFRRA	240
Db	237	SMSLGVLIIYLIVHGMPFPGDGHKTLVKQISNGAYRBPKPSPACGLIRLMLVNPFRRA	296
Qy	241	TLEDVYASHMWNVN	253
Db	297	TLEDVYASHMWNVN	309

```

RESULT 2
US-10-343-514-103
; Sequence 103, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBREYRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DDA-ORUC/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 103
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-514-103

```

Query Match	98.7%	Score 1325;	DB 16;	Length 251;
Best Local Similarity	100.0%	Pred. No. 1e-99;		
Matches 251; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	YEFLETLCKGYGKYKAKARESSGRVVAIKSIRKDKIKDDODLHRIIRREIIMSSLNPHI	60	
DB	1	YEFLETLCKGYGKYKAKARESSGRVVAIKSIRKDKIKDDODLHRIIRREIIMSSLNPHI	60	
QY	61	IAIHVEFENSSKIVIMEYASRGDLVDYIISERPRLSERDARHFRQIVALHYCHONGIV	120	
DB	61	IAIHVEFENSSKIVIMEYASRGDLVDYIISERPRLSERDARHFRQIVALHYCHONGIV	120	
QY	121	HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTGCSPLYASPEIYNGKPYGPEVD	180	
DB	121	HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTGCSPLYASPEIYNGKPYGPEVD	180	
QY	181	SMSSGLVLLIYLVHGMPFDGDQHKTLVKOISNGAYREPKPSDCAGLIWLMLVNPFTTRA	240	
DB	181	SMSSGLVLLIYLVHGMPFDGDQHKTLVKOISNGAYREPKPSDCAGLIWLMLVNPFTTRA	240	
QY	241	TLIEDVASHMWV	251	
DB	241	TLIEDVASHMWV	251	

```

RESULT 3
US-10-343-514-41
; Sequence 41, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.

```

```

: APPLICANT LEFEBVRE, Diana L.
: TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
: FILE REFERENCE: DPA-DR02/PCT
: CURRENT APPLICATION NUMBER: US/10/343,514
: PRIORITY FILING DATE: 2003-01-31
: PRIOR APPLICATION NUMBER: PCT/CA01/01109
: PRIOR FILING DATE: 2001-08-02
: PRIOR APPLICATION NUMBER: US 60/222,650
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 60/274,613
: PRIOR FILING DATE: 2001-03-12
: PRIOR APPLICATION NUMBER: CA 2,340,780
: PRIOR FILING DATE: 2001-03-28
: NUMBER OF SEQ ID NOS: 109
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 41
: LENGTH: 630
: TYPE: PRT
: ORGANISM: RAT
: US-10-343-514-41

```

Query Match	97.8%;	Score 1131;	DB 16;	Length 630;
Best Local Similarity	98.0%;	Pred. No. 2.78-98;		
Matches	248;	Conservative	3;	Mismatches 2;
			Indels	0;
			Gaps	0;
QY	1	YEFLETGKGTYGKVKAKARSSGRLVAIKSRKOKIKDEODLLHRRRIETMSSLNPHI	60	
Db	57	YEFLETGKGTYGKVKAKARSSGRLVAIKSRKOKIKDEODLLHRRRIETMSSLNPHI	116	
QY	61	IAIHVEVENSKIYIVMEYASRGDLVYISRRPLTSERDARHFFROIVSALHYCHONGIV	120	
Db	117	IAIHVEVENSKIYIVMEYASRGDLVYISRRPLNRDARHFFRQIVSALHYCHONGIV	176	
QY	121	HRDILKENIILLDANGNKIADFGSLNLYHGKPLQTFCSGSLVSPETVYNKKPPVGGPVD	180	
Db	177	HRDILKENIILLDANGNKIADFGSLNLYHGKPLQTFCSGSLVSPETVYNKKPPVGGPVD	236	
QY	181	SWSLGVLVLLVHGTMFPDGDHKTILVKQISNGAYRPPKPSDAGLIRWLMTNPTERRA	240	
Db	237	SWSLGVLVLLVHGTMFPDGDHKTILVKQISNGAYRPPKPSDAGLIRWLMTNPTERRA	296	
QY	241	TLEDVASHMMVNM 253		
Db	297	TLEDVASHMMVNM 309		

```

RESULT 4
US-10-343-514-50
; Sequence 50, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 251
; TYPE: PRT
; ORGANISM: RAT
US-10-343-514-50

```

Query Match	96.6%	Score 1396;	DB 16;	length 251;
Best Local Similarity	98.0%	Pred. No. 2,3e-97;		
Matches	246;	Conservative	3;	Mismatches 2;
		Indels	0;	Gaps 0
QY	1	YEFLETLGKGYGKYKKARSSGRLVAIKSIKDKIKDQODLHTRREIEIMSSLNHPHI	60	
Db	1	YEFLETLGKGYGKYKKARSSGRLVAIKSIKDKIKDQODLHTRREIEIMSSLNHPHI	60	
QY	61	IAIHEVFENSSKIIVIMVYASRGDLVDYISERPRLSERDARHFFQIVSALHYCHONGIV	120	
Db	61	IAIHEVFENSSKIIVIMVYASRGDLVDYISERPRLSERDARHFFQIVSALHYCHONGIV	120	
QY	121	HRDLKLENIILLDANENIKIADPGLSNLYHKGFLOTFCGSPLYASPELYNGKPYGPEVD	180	
Db	121	HRDLKLENIILLDASENIKIADPGLSNLYHKGFLOTFCGSPLYASPELYNGKPYGPEVD	180	
QY	181	SMSLGVLIIYLIVHGMPDPDGDHKLTVQIISGAVREPSPSDAGCLRIMLLMNPTRRA	240	
Db	181	SMSLGVLIIYLIVHGMPDPDGDHKLTVQIISGAVREPSPSDAGCLRIMLLMNPTRRA	240	
QY	241	TLIEDVASHMV 251		
Db	241	TLIEDVASHMV 251		

```

RESULT 5
US-09-963-159-2
; Sequence 2, Application US/0963159
; Patent No. US20020077312A1
; GENERAL INFORMATION:
; APPLICANT: Curtie, Rory A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 5700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 101/47-50U1
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-963-159-2

```

Query Match	96.3%;	Score 1293;	DB 9;	Length 628;
Best Local Similarity	95.7%;	Pred. No. 1,le-96;		
Matches	242;	Conservative	6;	Mismatches 5;
			Indels	0;
			Gaps	0;
QY	1	YEFLETGKGYGKYKARBSGRLVAIKSIRKDKIKOBODILHIRREIETMSSLNPHI	60	
Db	53	YEFLETGKGYGKYKARBSGRLVAIKSIRKDKIKOBODILHIRREIETMSSLNPHI	112	
QY	61	IAIHVFENSSKIVIVMEYASRGDIYDYSERPRKSERDARHFPFOIYSALFYCHONGIV	120	
Db	113	IAIHVFENSSKIVIVMEYASRGDIYDYSERQGLSERDARHFPFOIYSAYHYCHONRVV	172	
QY	121	HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLOTFCGSPLYASPEIYNGPYGPEVD	180	
Db	173	HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLOTFCGSPLYASPEIYNGPYGPEVD	232	
QY	181	SMSLGLVLLIYLIVHGMFPFGDGHKTLVQISNGARPPKPSDAGCLRMLLMVNPTRRA	240	
Db	233	SMSLGLVLLIYLIVHGMFPFGDGHKTLVQISNGARPPKPSDAGCLRMLLMVNPTRRA	292	
QY	241	TLEDVYASHMWNW	253	
Db	293	TLEDVYASHMWNW	305	

RESULT 6

```

US-10-423-543-44
; Sequence 44, Application US/10/423543
; Publication No. US20040058355A1
GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Liebermann, Rosana K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Kory A.J.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Miyoung
; APPLICANT: Williamson, Mark J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 21910, 56633, 55053, 2504, 15977,
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
; TITLE OF INVENTION: 18610, 33217, 21967, 11983, m1983, 38555 OR 593 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP103-0230MNM
; CURRENT APPLICATION NUMBER: US/10/423,543
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 10/278,036
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 09/711,216
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/205,447
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 10/012,055
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/248,325
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 10/003,690
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/248,893
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 10/217,168
; PRIOR FILING DATE: 2002-08-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-423-543-44

```

[illegible]

```
QY      241 TLEDVASHMWVM 253
Db      293 TLEDVASHMWVM 305

RESULT 7
US-10-618-941-77
; Sequence 77, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-77

Query Match      96.3%; Score 1293; DB 16; Length 672;
Best Local Similarity 95.7%; Pred. No. 1.2e-96;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      1 YEFLETGKGYGKVKKARESSGRVVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 60
Db      97 YEFLETGKGYGKVKKARESSGRVVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 156

QY      61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db      157 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERARHFFROIYSALHYCHONGIV 216

QY      121 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
Db      217 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 276

QY      181 SMSLGVLLYIVHGTMPPFGQDHKTIVKQISNGAYREPPKPSDACGLIRMLMVNPTTRA 240
Db      277 SMSLGVLLYIVHGTMPPFGQDHKTIVKQISNGAYREPPKPSDACGLIRMLMVNPTTRA 336

QY      241 TLEDVASHMWVM 253
Db      337 TLEDVASHMWVM 349

RESULT 8
US-10-311-034-12
; Sequence 12, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROMSKI, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAMLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
```

```
; APPLICANT: BURROD, Neil
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAPALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: DING, Li
; APPLICANT: GREYER, Megan
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: BATRA, Sajeiv
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
; CURRENT APPLICATION NUMBER: US/10/311,034
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372,
; 60/228,056
; PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
; 25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CD1
US-10-311-034-12

Query Match      95.8%; Score 1286; DB 15; Length 594;
Best Local Similarity 95.3%; Pred. No. 3.9e-96;
Matches 241; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      1 YEFLETGKGYGKVKKARESSGRVVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 60
Db      19 YEFLETGKGYGKVKKARESSGRVVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 78

QY      61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db      79 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERARHFFROIYSALHYCHONGIV 138

QY      121 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
Db      139 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 198

QY      181 SMSLGVLLYIVHGTMPPFGQDHKTIVKQISNGAYREPPKPSDACGLIRMLMVNPTTRA 240
Db      199 SMSLGVLLYIVHGTMPPFGQDHKTIVKQISNGAYREPPKPSDACGLIRMLMVNPTTRA 258

QY      241 TLEDVASHMWVM 253
Db      259 TLEDVASHMWVM 271

RESULT 9
US-10-343-514-101
; Sequence 101, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFSBARE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
```

FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 101
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-10-343-514-101

Query Match 95.1%; Score 1276; DB 16; Length 251;
Best Local Similarity 95.6%; Pred. No. 9.8e-96;
Matches 240; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 60
DB 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 60
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPLSERDARHPROQVSAHYCHONGIV 120
DB 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPLSERDARHPROQVSAHYCHONGIV 120
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLYASPEIYVNGKPYGPEVD 180
DB 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLYASPEIYVNGKPYGPEVD 180
QY 181 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 181 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
QY 241 TLEDVASHMWV 251
DB 241 TLEDVASHMWV 251

RESULT 10
US-10-322-281-23
Sequence 23; Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 640
TYPE: PRT
ORGANISM: Mus musculus
US-10-322-281-23

Query Match 94.9%; Score 1273; DB 16; Length 640;
Best Local Similarity 97.2%; Pred. No. 4.8e-95;
Matches 246; Conservative 0; Mismatches 1; Indels 6; Gaps 3;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 60
DB 72 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 131
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPLSERDARHPROQVSAHYCHONGIV 120

DB 132 IAIHEVEFENSSKIVIMEVYASRGDLVDYIS--PRLS--DARHPROQVSAHYCHONGIV 187
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLYASPEIYVNGKPYGPEVD 180
DB 188 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLY--DEIVNGKPYGPEVD 245
QY 181 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 246 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 305
QY 241 TLEDVASHMWV 253
DB 306 TLEDVASHMWV 318

RESULT 11
US-10-322-281-26
Sequence 26; Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 616
TYPE: PRT
ORGANISM: Homo sapiens
US-10-322-281-26

Query Match 91.7%; Score 1230; DB 16; Length 616;
Best Local Similarity 93.3%; Pred. No. 1.4e-91;
Matches 236; Conservative 6; Mismatches 5; Indels 6; Gaps 3;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 60
DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNHPHI 112
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPLSERDARHPROQVSAHYCHONGIV 120
DB 113 IAIHEVEFENSSKIVIMEVYASRGDLVDYISQ--QLSE--ARHPROQVSAHYCHONGIV 168
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLYASPEIYVNGKPYGPEVD 180
DB 169 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCCGSPLY--DEIVNGKPYGPEVD 226
QY 181 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 227 SMSLGLVLYLVHGTMPEFGQDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 286
QY 241 TLEDVASHMWV 253
DB 287 TLEDVASHMWV 299

RESULT 12
US-09-780-949-2
Sequence 2; Application US/09780949
Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Welch, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690

;; PRIOR FILING DATE: 2000-02-09
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 661
;; TYPE: PRF
;; ORGANISM: H. sapiens
US-09-780-949-2

Query Match 84.1%; Score 1128.5; DB 9; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVKKAKES-SGRVVAIKSIRKDKIKEDQDLHTRREIETMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYKVRATERFSGRVVAIKSIRKDKIKEDQDVHTRREIETMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSCKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASRGELVDYISERRRISERETRHFRQIVSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 VHRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPUYASPEIVNGRPYVGP 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 175 VHRDLKLENTLLDNCNIKIADFGLSNLVQKDFLOTFCGSPUYASPEIVNGRPYGP 234
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 180 DSWSLGVLLYILVHGTMPPDGDHKTIVKQISNGATREPKPSDAGCLIRMLMNPTR 239
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 235 DSWALGVLLYTLVYGTMPDGFPHKNIIRQISSGEYREPTQSDARGLIRMLMNPDRR 294
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 ATLEDVASHMWVM 253
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 295 ATLEDIANHWVM 308
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
US-09-780-949-6
;; Sequence 6, Application US/09780949
;; Patent No. US20020006618x1
;; GENERAL INFORMATION:
;; APPLICANT: Kapeller-Liberman, Rosana
;; APPLICANT: Welch, Nadine S.
;; APPLICANT: Galvin, Katherine M.
;; TITLE OF INVENTION: Methods for Using 20893, a Human Protein
;; TITLE OF INVENTION: Kinase
;; FILE REFERENCE: 035800/209015
;; CURRENT APPLICATION NUMBER: US/09/780,949
;; CURRENT FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: US 60/181,690
;; PRIOR FILING DATE: 2000-02-09
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 661
;; TYPE: PRF
;; ORGANISM: H. sapiens
US-09-780-949-6

Query Match 84.1%; Score 1128.5; DB 9; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVKKAKES-SGRVVAIKSIRKDKIKEDQDLHTRREIETMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYKVRATERFSGRVVAIKSIRKDKIKEDQDVHTRREIETMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSCKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASRGELVDYISERRRISERETRHFRQIVSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 VHRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPUYASPEIVNGRPYVGP 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 175 VHRDLKLENTLLDNCNIKIADFGLSNLVQKDFLOTFCGSPUYASPEIVNGRPYGP 234
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 180 DSWSLGVLLYILVHGTMPPDGDHKTIVKQISNGATREPKPSDAGCLIRMLMNPTR 239
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 235 DSWALGVLLYTLVYGTMPDGFPHKNIIRQISSGEYREPTQSDARGLIRMLMNPDRR 294
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 ATLEDVASHMWVM 253
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 295 ATLEDIANHWVM 308
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
US-10-354-358-82
;; Sequence 82, Application US/10354358
;; Publication No. US20030157082A1
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc
;; APPLICANT: Hunter, John Joseph
;; APPLICANT: MacBeth, Kyle J.
;; APPLICANT: Teal, Fong-Ying
;; APPLICANT: Leeson, Andrea
;; APPLICANT: Lightcap, Eric S.
;; APPLICANT: Williamson, Mark
;; APPLICANT: Rudolph-Owen, Laura A.
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
;; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
;; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
;; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
;; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
;; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
;; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
;; FILE REFERENCE: MP102-0201RNMNTM
;; CURRENT APPLICATION NUMBER: US/10/354,358
;; PRIOR FILING DATE: 2003-01-30
;; PRIOR APPLICATION NUMBER: US 60/353,600
;; PRIOR FILING DATE: 2002-01-31
;; PRIOR APPLICATION NUMBER: US 60/364,517
;; PRIOR FILING DATE: 2002-03-15
;; PRIOR APPLICATION NUMBER: US 60/371,075
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: US 60/371,507
;; PRIOR FILING DATE: 2002-04-10
;; PRIOR APPLICATION NUMBER: US 60/372,984
;; PRIOR FILING DATE: 2002-04-16
;; PRIOR APPLICATION NUMBER: US 60/374,194
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: US 60/382,995
;; PRIOR FILING DATE: 2002-05-24
;; PRIOR APPLICATION NUMBER: US 60/385,023
;; PRIOR FILING DATE: 2002-05-31
;; PRIOR APPLICATION NUMBER: US 60/388,853
;; PRIOR FILING DATE: 2002-06-14
;; PRIOR APPLICATION NUMBER: US 60/389,395
;; PRIOR FILING DATE: 2002-06-17
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 122
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 82
;; LENGTH: 661
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-354-358-82

Query Match 84.1%; Score 1128.5; DB 14; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVKKAKES-SGRVVAIKSIRKDKIKEDQDLHTRREIETMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYKVRATERFSGRVVAIKSIRKDKIKEDQDVHTRREIETMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSCKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASRGELVDYISERRRISERETRHFRQIVSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 120 VHRDLKLENILLDANGNIIKIADFGISNLVHKGKFLQTFGSPLYASPEIYNGKPYVGPV 179
 DB 175 VHRDLKLENILLDANGNIIKIADFGISNLVQKQFLQTFGSPLYASPEIYNGRPRGPV 234
 QY 180 DMSVSGVLLIYIVHGTMPDPGQDHKTLYVQISNGAYREPPKPSDACGLIRMLMNPTR 239
 DB 235 DSWALGVLLIYIVHGTMPDPGFDHKNLIRQISSGAYREPTQPSDARGLIRMLMNPDRR 294
 QY 240 ATLEDVASHMWV 253
 DB 295 ATLEDIANHWMV 308

RESULT 15
 US-10-343-514-102
 ; Sequence 102, Application US/10343514
 ; Publication No. US20040132025A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUCKER, Daniel J.
 ; APPLICANT: ROSEN, Cheryl F.
 ; APPLICANT: LEFEBVRE, Diana L.
 ; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
 ; FILE REFERENCE: DPA-DRUC2/ECT
 ; CURRENT APPLICATION NUMBER: US/10/343,514
 ; CURRENT FILING DATE: 2003-01-31
 ; PRIOR APPLICATION NUMBER: PCT/CA01/01109
 ; PRIOR FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: US 60/222,650
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/274,613
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: CA 2,340,780
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 102
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-343-514-102

Query Match 82.8%; Score 1111.5; DB 16; Length 252;
 Best Local Similarity 82.1%; Pred. No. 2.3e-82;
 Matches 207; Conservative 24; Mismatches 20; Indels 1; Gaps 1;
 QY 1 YEPLETGKGTGKVKKARE-SGRVLAISKIRKDKIKDQDLIARRREIEMSLNHPH 59
 DB 1 YELQETLGTGKGTGKVKRATERPSGRVVAISKIRKDKIKDQDMVHARRREIEMSLNHPH 60
 QY 60 IIAIHEVFENSSKIVIVMEYASRGDLVDYISERPRLSERDARHPFQIYSALHYCHONGI 119
 DB 61 IISIEVFENKDKIVIMBYASKGLYDIISERRRISERETRRHFRQIVSAVHYCHONGV 120
 QY 120 VHRDLKLENILLDANGNIIKIADFGISNLVHKGKFLQTFGSPLYASPEIYNGKPYVGPV 179
 DB 121 VHRDLKLENILLDANGNIIKIADFGISNLVQKQFLQTFGSPLYASPEIYNGRPRGPV 180
 QY 180 DMSVSGVLLIYIVHGTMPDPGQDHKTLYVQISNGAYREPPKPSDACGLIRMLMNPTR 239
 DB 181 DSWALGVLLIYIVHGTMPDPGFDHKNLIRQISSGAYREPTQPSDARGLIRMLMNPDRR 294
 QY 240 ATLEDVASHMWV 251
 DB 241 ATLEDIANHWMV 252

Search completed: May 11, 2005, 14:42:21
 Job time : 53.0882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:09:18 ; Search time 66.3982 Seconds
(without alignments)
1473.690 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309
Perfect score: 1342
Sequence: 1 YEFLETLGKGYGVKVKARE.....VNPTRRATLEADVASHMWVNW 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	4	AAB50056 Murine ly
2	1342	100.0	652	8	ABO84756 Murine ca
3	1313	97.8	630	5	AAE19885 Rat SNF1/
4	1293	96.3	628	4	AAAM93360 Human pol
5	1293	96.3	628	4	ABU53319 Human cel
6	1293	96.3	628	5	ABP69116 Human pro
7	1293	96.3	628	5	AAU79652 Human pro
8	1293	96.3	628	7	ADP76965 Novel hum
9	1293	96.3	628	8	ADL30886 Human pro
10	1293	96.3	628	8	ADL25362 Human SNA
11	1293	96.3	628	8	ADL14161 Novel hum
12	1293	96.3	628	8	ABO84757 Human PRO
13	1293	96.3	628	8	ADL20172 Human can
14	1293	96.3	628	8	ADJ96650 Human cal
15	1286	95.8	594	5	AAE16266 Human kin
16	1279.5	95.3	611	6	ABP6085 Human pro
17	1268.5	94.5	629	4	AAE5632 Novel pro
18	1268.5	94.5	629	8	ADL29239 Human MAR
19	1128.5	84.1	660	7	ADN95766 Human BEC
20	1128.5	84.1	661	4	AAE07847 Human pro
21	1128.5	84.1	661	4	AAE07846 Human pro
22	1128.5	84.1	661	7	ADP38421 Human pro
23	1128.5	84.1	661	8	ADU75331 Mar ker ge
24	1128.5	84.1	661	8	ADL25353 Human ARK
25	1128.5	84.1	661	8	ADQ19734 Human sof

26	1085	80.8	534	4	AAB71959 Aab71959 Human TGF
27	822.5	61.3	434	4	ABE67451 ABE67451 Drosophila
28	814	60.7	530	4	AAE00668 Human pro
29	665.5	49.6	508	5	ABG79179 Human ser
30	665.5	49.6	520	4	AAE5629 Human pro
31	665.5	49.6	520	8	ADL29235 Mouse MAR
32	665.5	49.6	688	5	ABE04434 Human neu
33	665.5	49.6	689	7	ADP74129 Human neu
34	665.5	49.6	752	4	AAE11782 Human kin
35	665.5	49.6	752	4	AAE11782 Human kin
36	665.5	49.6	752	5	ABE04433 Human neu
37	665.5	49.6	752	6	ABG73794 Human MAR
38	665.5	49.6	752	8	ADJ32125 Human pro
39	665.5	49.6	769	5	ABP62966 Human pro
40	661.5	49.3	752	8	ADJ96622 Human cal
41	660.5	49.2	660	5	ABE04432 Murine ne
42	660.5	49.2	688	6	AAE33554 Human mic
43	660.5	49.2	688	6	AAE33555 Human mic
44	660.5	49.2	688	7	ADG91726 Human mic
45	660.5	49.2	688	7	ADG91727 Human mic

ALIGNMENTS

RESULT 1	
AB50056	
ID	AB50056 standard; protein; 631 AA.
XX	
AC	AB50056;
XX	
DT	19-MAR-2001 (first entry)
XX	
DE	Murine Lymph node Stromal cell kinase 1.
XX	
KW	Murine; Lymph node Stromal cell kinase; MLK-1; autoimmune disorder;
KW	wound healing; peridontal disease; inflammatory disease; tumour;
XX	
XX	infection; allergy.
XX	
OS	Mus musculus.
XX	
PN	W0200073468-A1.
XX	
PD	07-DEC-2000.
XX	
PP	26-MAY-2000; 2000WO-US014696.
XX	
PR	28-MAY-1999; 99US-0136781P.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Bird TA, Virca GD, Martin U, Anderson DM;
XX	
DR	WPI; 2001-061546/07.
XX	
DR	N-PSDB; AAC0433.
XX	
PT	Novel murine and human kinase nucleic acids useful for treating
PT	inflammations, infections, tumors, allergies, autoimmune diseases, and
PT	for stimulating or suppressing immune responses.
XX	
PS	Claim 10; Page 94-96; 106pp; English.
XX	
CC	The present sequence is Murine Lymph node Stromal cell kinase 1 (MLK-1).
CC	This protein is useful for treating a variety of disorders listed in the
CC	disclosure of the specification, including autoimmune disorders, allergic
CC	reactions, myeloid or lymphoid cell deficiencies, wound healing, and
CC	tissue repair and replacement, burns, incisions and ulcers, peridontal
CC	disease, inflammatory diseases, tumours and bacterial, viral or fungal
CC	infection
XX	
XX	Sequence 631 AA;
SQ	
Query Match	100.0%; Score 1342; DB 4; Length 631;

Best Local Similarity 100.0%; Pred. No. 2.7e-138; Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLLHIREIEMSLNHPH 60
Db 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLLHIREIEMSLNHPH 116
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
QY 181 SMSLGLVLYILVHGTMPPDGDHKTLYVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
Db 237 SMSLGLVLYILVHGTMPPDGDHKTLYVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 296
QY 241 TLEDVASHMWVNW 253
Db 297 TLEDVASHMWVNW 309
```

RESULT 2

AB084756
ID AB084756 standard; protein; 652 AA.

AC AB084756;

DT 18-NOV-2004 (first entry)

DE Murine cancer-associated protein (CAP) M07-004.

KW Mouse; cancer-associated protein; CAP; cancer; cytostatic.

OS Mus musculus.

PN W02004058146-A2.

PD 15-JUL-2004.

PF 15-DEC-2003; 2003WO-US040081.

PK 17-DEC-2002; 2002US-00322281.

PA (SAGR-) SAGRES DISCOVERY INC.

PI Morris DW, Malandro MS;

DR WPI; 2004-499109/47.

DR N-PSDB; ABD33083.

PT Novel human cancer associated protein encoded within open reading frame

PT of cancer associated gene, useful as targets for diagnosing cancer.

PS Disclosure; SEQ ID NO 23; 182pp; English.

CC The invention relates to cancer-associated proteins (CAP) and the cancer-associated (CA) nucleic acids encoding them. The invention also relates to a method for treating cancers involving administering to a patient an inhibitor of CAP, and a method of screening for anticancer activity in a potential drug involving providing a cell that expresses a CA gene, contacting a tissue sample derived from a cancer cell with an anticancer drug candidate and monitoring the effect of the anticancer drug candidate on expression of the CA gene. The CAP proteins are useful for detecting cancer associated with expression of a CAP protein in a test cell sample and for screening for a bioactive agent capable of modulating the activity of a CAP protein. The CA nucleic acids are useful for diagnosing cancer, involving determining the expression of a CA nucleic acid in a tissue. This sequence represents a murine CAP of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 652 AA;

Query Match 100.0%; Score 1342; DB 8; Length 652;

Best Local Similarity 100.0%; Pred. No. 2.8e-138; Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLLHIREIEMSLNHPH 60
Db 78 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLLHIREIEMSLNHPH 137
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db 138 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 197
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
Db 198 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 257
QY 181 SMSLGLVLYILVHGTMPPDGDHKTLYVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
Db 258 SMSLGLVLYILVHGTMPPDGDHKTLYVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 317
QY 241 TLEDVASHMWVNW 253
Db 318 TLEDVASHMWVNW 330
```

RESULT 3

AAE19885
ID AAE19885 standard; protein; 630 AA.

AC AAE19885;

DT 18-JUN-2002 (first entry)

DE Rat SNF1/AMPK-Related Kinase (SNARK) protein.

KW Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;

KW glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;

KW hyperglycaemic; drug screening; hypoglycaemia.

OS Rattus sp.

XX XX

FH Key

FT Binding-site Location/Qualifiers

FT Domain /note= "Protein kinase ATP-binding region signature"

FT Active-site /note= "Serine/threonine kinase catalytic domain"

FT Domain /note= "Serine/threonine protein kinase active-site"

FT Domain /note= "Serine/threonine kinase catalytic domain"

FT Domain /note= "Serine/threonine kinase catalytic domain"

FT Domain /note= "Serine/threonine kinase catalytic domain"

FT Domain /note= "Serine/threonine kinase catalytic domain"

FT Domain /note= "Serine/threonine kinase catalytic domain"

PN W0200212456-A2.

PD 14-FEB-2002.

XX XX

PF 02-AUG-2001; 2001MO-CA001109.
XX
XX 03-AUG-2000; 2000US-0222650P.
PR 12-MAR-2001; 2001US-0274613P.
PR 26-MAR-2001; 2001CA-02340783.
XX
XX (ONCO-) 1149336 ONTARIO INC.
XX
XX Drucker DJ, Rosen CF, Lefebvre DL,
XX
XX WPI; 2002-241747/29.
DR N-PSDB; AAD31710.
XX
XX AMPK (AMP-activated protein kinase)-related kinase, designated SNARK
PT polypeptides and polynucleotides, useful for treating or preventing
PT diabetes, or other disorders of lipoprotein production leading to
XX increased levels of cholesterol.
XX
XX Claim 1; Fig 2; 94pp; English.
XX
XX The invention relates to an AMPK (AMP-activated protein kinase)-related
CC kinase, designated SNARK polypeptides and polynucleotides. SNARK
CC (SNIT/AMP-activated protein kinase) is involved in stress response to
CC glucose deprivation. The polynucleotides are useful for expressing SNARK
CC protein in isolated form or as a protein conjugate. Activation of SNARK
CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
CC in other cell types such as heart and skeletal muscles, as well as
CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
CC to have insulin-like effects that would enhance the disposal of glucose
CC into muscle and reduce plasma glucose for the treatment of diabetes and
CC some type of disorders of lipoprotein production leading to increased
CC levels of cholesterol or triglycerides. SNARK or its variants may be
CC administered to a subject to treat or prevent a disease associated with
CC decreased expression of SNARK, such as diabetes. SNARK antibodies are
CC used to modulate SNARK activity either in vivo for therapeutic purposes,
CC or in vitro, for drug screening and related investigational purposes.
CC SNARK antagonists may be administered to increase fuel production,
CC decrease glucose uptake and increase levels of blood glucose in a patient
CC suffering from hypoglycaemia. The present sequence is rat SNARK protein
XX
XX Sequence 630 AA;
SQ

Query Match 97.8%; Score 1313; DB 5; Length 630;
Best Local Similarity 98.0%; Pred. No. 4.3e-135;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIETIMSSLNHPHI 60
DB 57 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIETIMSSLNHPHI 116
QY 61 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAALHYCHONGIV 120
DB 117 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAALHYCHONGIV 176
QY 121 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 180
DB 177 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 236
QY 181 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYREPPKPSDAGLIRWLLVNPTRRA 240
DB 237 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYREPPKPSDAGLIRWLLVNPTRRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309

RESULT 4
ID AAM93360 standard; protein; 628 AA.
XX
XX AAM93360;
XX

DT 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 2919.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX PD 05-SEP-2001.
XX
XX PF 07-JUL-2000; 2000EP-00114089.
XX
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Nishikawa T, Isogai T, Hayaeshi K, Ishii S, Kawai Y,
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94280.
XX
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX PS Claim 8; SEQ ID NO 2919; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
XX
XX format directly from EPO
SQ

Query Match 96.3%; Score 1293; DB 4; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIETIMSSLNHPHI 60
DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIETIMSSLNHPHI 112
QY 61 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAALHYCHONGIV 120
DB 113 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSAALHYCHONGIV 172
QY 121 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 180
DB 173 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 232
QY 181 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYREPPKPSDAGLIRWLLVNPTRRA 240
DB 233 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYREPPKPSDAGLIRWLLVNPTRRA 292
QY 241 TLEDVASHMWVNM 253
DB 293 TLEDVASHMWVNM 305

RESULT 5
ABUS3319

ID ABU5319 standard; protein; 628 AA.
XX AC ABU5319;
XX DT 14-ADR-2003 (first entry)
XX DE Human cell cycle-associated protein from DKFzphes3_7j3.
XX KM Human; gene therapy; vaccine; disease treatment; detection.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-IB001496.
XX PR 18-AUG-1999; 99US-0149499P.
XX PR 28-SEP-1999; 99US-0156503P.
XX PA (GENU-) GERMAN HUMAN GENOME PROJECT.
XX PI Wiemann S;
XX DR WPI; 2001-327840/34.
XX DR N-PSDB; ABX71420.
XX PT Nucleic acids having the sequences of clones isolated from libraries of
XX PS different human tissues, useful in recombinant DNA methodologies.
XX Claim 21; Page 943; 1095pp; English.
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a
CC polypeptide described in the disclosure of the invention
XX
XX Sequence 628 AA;
SQ
Query Match 96.3%; Score 1293; DB 4; Length 628;
Best Local Similarity 95.7%; Pred. No. 6,8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETLGKGYGKVKKARESSGRVLAISIRKDKIKEDODLHIREIFMISLNHPHI 60
DB 53 YEFLETLGKGYGKVKKARESSGRVLAISIRKDKIKEDODLHIREIFMISLNHPHI 112
QY 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIIVSAHYCHONGIV 120
DB 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERQULSEREAKHFFROIIVSAHYCHONGIV 172
QY 121 HEDLKIENILLDANGNIKIADFGLSNLYHNGKFLQTFCCGSPLYASPEIYVNGKPYGPEVD 180
DB 173 HEDLKIENILLDANGNIKIADFGLSNLYHNGKFLQTFCCGSPLYASPEIYVNGKPYGPEVD 232
QY 181 SMSLGLVLYLVHGMTPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLVMVNPTRRA 240
DB 233 SMSLGLVLYLVHGMTPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLVMVNPTRRA 292
QY 241 TLEDVASHMMVMW 253
DB 293 TLEDVASHMMVMW 305

ID ABP69116 standard; protein; 628 AA.
XX AC ABP69116;
XX DT 20-JAN-2003 (first entry)
XX DE Human polypeptide SEQ ID NO 1163.
XX KM Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX KM cell-proliferative disorder; neurodegenerative disease; bacterial;
XX KM Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX KM arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
XX KM antiparkinsonian; antididiabetic; immunosuppressive; dermatological;
XX KM haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
XX KM antiarthritic.
XX OS Homo sapiens.
XX PN WO200270539-A2.
XX PD 12-SEP-2002.
XX PF 05-MAR-2002; 2002WO-US005095.
XX PR 05-MAR-2001; 2001US-00799451.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX DR WPI; 2002-759812/82.
XX DR N-PSDB; ABZ11333.
XX PT New polynucleotides comprising sequences assembled from expressed
XX PS sequence tags (ESTs), useful for treating cell-proliferative,
XX PS neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
XX PS or coagulation disorders.
XX Claim 9; SEQ ID NO 1163; 1012pp + Sequence Listing; English.
CC The invention relates to an isolated polynucleotide (1) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pot_sequences
XX
XX Sequence 628 AA;
SQ
Query Match 96.3%; Score 1293; DB 5; Length 628;
Best Local Similarity 95.7%; Pred. No. 6,8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETLGKGYGKVKKARESSGRVLAISIRKDKIKEDODLHIREIFMISLNHPHI 60
DB 53 YEFLETLGKGYGKVKKARESSGRVLAISIRKDKIKEDODLHIREIFMISLNHPHI 112
QY 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIIVSAHYCHONGIV 120
DB 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERQULSEREAKHFFROIIVSAHYCHONGIV 172

QY 121 HRDLKLENTLLDANGNIKIADFGSLNLYHKGFLOTFCCSPLYASPEIYNGKPYGPEVD 180
CC |||||
XX |||||
DB 173 HRDLKLENTLLDANGNIKIADFGSLNLYHQGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGVLYLYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVPTRRA 240
DB 233 SMSLGVLYLYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVPTRRA 292
QY 241 TLEDVASHMWVMW 253
DB 293 TLEDVASHMWVMW 305

RESULT 7
AAU79652
ID AAU79652 standard; protein; 628 AA.
XX
AC AAU79652;
DT 02-JUL-2002 (first entry)
XX
DE Human protein kinase 3700.
XX
KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
KW cell proliferation disorder; cell differentiation disorder; carcinoma;
KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;
KW cytosolic; antiatherosclerotic; enzyme.
XX
OS Homo sapiens.
XX
PN WO200224921-A2.
XX
PD 28-MAR-2002.
XX
PF 25-SEP-2001; 2001WO-US030115.
XX
PR 25-SEP-2000; 2000US-0234922P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Galvin KM;
XX
DR MPI; 2002-352007/38.
DR N-PSDB; ABK14000.
XX
PT Use of modulators of activity of 3700 protein for making medicament for
PT e.g., modulating protein phosphorylation or cell signaling, or for
PT treating or preventing cellular proliferative and/or differentiative
PT disorders.
XX
PS Claim 19; Fig 1; 115pp; English.
XX
CC The present invention relates to the isolation of a novel human protein
CC kinase designated 3700, and the polynucleotide sequence encoding it. The
CC invention also describes the use of a modulator of the activity of
CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
CC composition for modulating the ability of a cell to phosphorylate an
CC amino acid residue of a substrate protein. Modulators of protein kinase
CC 3700 activity are useful for modulating protein phosphorylation, cell
CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
CC angiogenesis, tissue repair, tissue regeneration, establishment or
CC progression of atherosclerosis, and signalling across the blood-brain
CC barrier. The polynucleotide and polypeptide molecules for protein kinase
CC 3700 may be used as diagnostic targets and therapeutic agents for
CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
CC curing PK-related disorders and cellular proliferative and/or
CC differentiative disorders (e.g. haematopoietic neoplastic disorders,
CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
CC polynucleotide sequence can be used to express protein kinase 3700, to
CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for

CC tissue typing, in forensic biology, and as surrogate markers. The present
CC sequence represents human protein kinase 3700
XX
SQ Sequence 628 AA;
QY Query Match 96.3%; Score 1293; DB 5; Length 628;
DB Best Local Similarity 95.7%; Pred. No. 6,8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETGKGTGKVKAKRESSGRLVAIKSIRKDKIKDDOLLHRRRIEIMSLNPHI 60
DB 53 YEFLETGKGTGKVKAKRESSGRLVAIKSIRKDKIKDDOLLHRRRIEIMSLNPHI 112
QY 61 IAIHEVFENSSKIYIVMEYASRDLYDIYISRRPRLSERDARHFRQIVSALHYCHONGIV 120
DB 113 IAIHEVFENSSKIYIVMEYASRDLYDIYISRRQSLSERDARHFRQIVSAVHCQGNRVV 172
QY 121 HRDLKLENTLLDANGNIKIADFGSLNLYHKGFLOTFCCSPLYASPEIYNGKPYGPEVD 180
DB 173 HRDLKLENTLLDANGNIKIADFGSLNLYHQGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGVLYLYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVPTRRA 240
DB 233 SMSLGVLYLYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVPTRRA 292
QY 241 TLEDVASHMWVMW 253
DB 293 TLEDVASHMWVMW 305

RESULT 8
ADF76965
ID ADF76965 standard; protein; 628 AA.
XX
AC ADF76965;
DT 26-FEB-2004 (first entry)
XX
DE Novel human secreted and transmembrane protein Seqid 640.
XX
KW human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytosolic; chondrocyte; tumour.
XX
OS Homo sapiens.
XX
PN WO2003072035-A2.
XX
PD 04-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US005241.
XX
PR 22-FEB-2002; 2002US-0359461P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
PI Williams PM, Wood WT, Wu TD;
XX
DR MPI; 2003-721702/68.
DR N-PSDB; ADF76964.
XX
PT New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g., systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.
XX
PS Claim 10; SEQ ID NO 640; 918pp; English.
XX
CC This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles

CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC stimulation of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is the amino acid sequence of a human PRO protein of the
CC invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 7; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEODLHIREIEMSSLNHPH 60
DB 53 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEODLHIREIEMSSLNHPH 112
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERARHFFRQIVSAVHYCHONRV 172
QY 121 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
DB 173 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 233 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 292
QY 241 TLEDVASHMWVMV 253
DB 293 TLEDVASHMWVMV 305

RESULT 9
ADL30886
ID ADL30886 standard; protein; 628 AA.

AC ADL30886;
DT 20-MAY-2004 (first entry)

DE Human protein encoded by a full length cDNA clone Segid 2919.

KW human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method.

OS Homo sapiens.

XX EPI396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Oca T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
DR N-PSDB; ADL30885.

XX New oligonucleotide primers (830 CDNAs) useful for synthesizing full
PT length human cDNAs.

PS Example 1; SEQ ID NO 2919; 1340bp; English.

CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEODLHIREIEMSSLNHPH 60
DB 53 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEODLHIREIEMSSLNHPH 112
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERARHFFRQIVSAVHYCHONRV 172
QY 121 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
DB 173 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 233 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYREPPKPSDAGLIRMLMVNPTTRA 292
QY 241 TLEDVASHMWVMV 253
DB 293 TLEDVASHMWVMV 305

RESULT 10
ADL25362
ID ADL25362 standard; protein; 628 AA.

AC ADL25362;

DT 03-JUN-2004 (first entry)

DE Human SWARK, SEQ ID 12.

KW Cytostatic; Neuroprotective; Nootropic; Muscular; Gene therapy;
XX Adenosine Monophosphate activated Protein Kinase-related kinase 5;
XX AMPK-related kinase 5; ARK5; KIAA0537; stress resistance; tumour;
XX nervous disorder; muscle disorder; ataxia telangiectasia; SWARK.

OS Homo sapiens.

XX WO2004019994-A1.

XX 11-MAR-2004.

XX 19-AUG-2003; 2003WO-JP010435.

CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
CC 21967, 11983, 11983, 38555 or 593 nucleic acid molecule (1) comprising
CC any one of 40 nucleotide sequences (1). The nucleic acid molecules and
CC polypeptides are useful for diagnosing and treating a subject having a
CC disorder, or a subject at risk of developing a disorder, which is
CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, 11983,
CC 11983, 38555 or 593 activity, such as cellular proliferative and/or
CC differentiative disorders, brain disorders, platelet disorders, breast
CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
CC ovarian disorders, prostate disorders, cervical disorders, spleen
CC disorders, thymus disorders, thyroid disorders, testes disorders,
CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
CC skin (dermal) disorders, disorders associated with bone metabolism,
CC immune, e.g., inflammatory disorders, cardiovascular disorders,
CC endometrial cell disorders, liver disorders, viral diseases, pain
CC disorders, metabolic disorders, neurological or central nervous system
CC disorders, erythroid disorders, blood vessel disorders or angiogenic
CC disorders (all claimed), e.g., cancer, heart failure, hypertension,
CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
CC disease, psoriasis, or asthma. The nucleic acid molecules and
CC polypeptides are also useful as modulating agents in regulating a variety
CC of cellular processes, e.g., cell proliferation, differentiation, growth and
CC division. This is the amino acid sequence of a novel human protein of the
CC invention. Note: The sequences given in the specification are also
CC available in electronic format from
CC ftp.segdata.uspto.gov/Sequence.html?DocID=20040058355.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;

Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGYKVKKARSSGRVLAISKIRKDKDEODLHIREIEMSSLNHPH 60

DB 53 YEFLETLGKGTGYKVKKARSSGRVLAISKIRKDKDEODLHIREIEMSSLNHPH 112

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120

DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172

QY 121 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 180

DB 173 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 232

QY 181 SMSLGLVLYLVHGTMPPFGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMNPTRRA 240

DB 233 SMSLGLVLYLVHGTMPPFGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMNPTRRA 292

QY 241 TLEDVASHMWVNW 253

DB 293 TLEDVASHMWVNW 305

Db

RESULT 12
ADO20172 standard; protein; 628 AA.

XX ADO20172;

XX 12-AUG-2004 (first entry)

XX Human PRO polypeptide #540.

XX Human; PRO; immune related disorder; systemic lupus erythematosus;
XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
XX systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
XX autoimmune haemolytic anaemia; autoimmune chromocytopenia; thyroiditis;
XX diabetes mellitus; renal disease; demyelinating disease;
XX central nervous system; peripheral nervous system;
XX demyelinating polyneuropathy; Guillain-Barre syndrome;
XX chronic inflammatory demyelinating polyneuropathy.

XX OS Homo sapiens.
XX WO2004043361-A2.
XX 27-MAY-2004.
XX 06-NOV-2003; 2003WO-US035268.
XX 08-NOV-2002; 2002US-0425235P.
XX (GENTH) GENTECH INC.
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
XX Wood WJ, Wu TD;
XX WPI; 2004-420067/39.
XX N-PSDB; ADO20171.
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
XX treating an immune related disorder such as systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX spondyloarthropathy.
XX Claim 7; SEQ ID NO 1080; 1731pp; English.

XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;

Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGYKVKKARSSGRVLAISKIRKDKDEODLHIREIEMSSLNHPH 60

DB 53 YEFLETLGKGTGYKVKKARSSGRVLAISKIRKDKDEODLHIREIEMSSLNHPH 112

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120

DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172

QY 121 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 180

DB 173 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 232

QY 181 SMSLGLVLYLVHGTMPPFGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMNPTRRA 240

DB 233 SMSLGLVLYLVHGTMPPFGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMNPTRRA 292

QY 241 TLEDVASHMWVNW 253

DB 293 TLEDVASHMWVNW 305

Db

RESULT 13
AB084757 standard; protein; 628 AA.

XX AB084757;

XX 18-NOV-2004 (first entry)

XX Human cancer-associated protein (CAP) HP07-004.
DE Human; cancer-associated protein; CAP; cancer; cytostatic.
XX Homo sapiens.
XX MO2004058146-A2.
XX 15-JUL-2004.
XX 15-DEC-2003; 2003MO-US040081.
XX 17-DEC-2002; 2002US-00322281.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX WPI; 2004-499109/47.
XX N-PSDB; ABD33085.
XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX Claim 18; SEQ ID NO 26; 182bp; English.
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a human CAP of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 628 AA;
Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETLGKGTGKVKKARSSGRVAIKSIRKDKDEODLHIREIRIEMSLNPHI 60
DB 53 YEFLETLGKGTGKVKKARSSGRVAIKSIRKDKDEODLHIREIRIEMSLNPHI 112
QY 61 IAIHEVFENSSKIVIMEVYASRGDLVDYISERPRLSERDARHFRQIVSAHYCHONGIV 120
DB 113 IAIHEVFENSSKIVIMEVYASRGDLVDYISERQOLSREARHFRQIVSAHYCHONGIV 172
QY 121 HRDLKLENTLLDANGNINKIADFGLSNLYHGKFLQTFCCSPLVYASPEIVNGKPYGPEVD 180
DB 173 HRDLKLENTLLDANGNINKIADFGLSNLYHGKFLQTFCCSPLVYASPEIVNGKPYGPEVD 232
QY 181 SWSLGVLLYLVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRWLWVNPTRRA 240
DB 233 SWSLGVLLYLVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRWLWVNPTRRA 292
QY 241 TLEDVASHMWVNM 253
DB 293 TLEDVASHMWVNM 305
RESULT 14
ADJ96620

ID ADJ96620 standard; protein; 672 AA.
XX ADJ96620;
AC XX
XX 06-MAY-2004 (first entry)
DT XX
XX Human calcium/calmodulin-dependent protein kinase Nkak2 protein Seqid 77.
DE XX
XX kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
XX PKR; STK; gene therapy; cancer; immune-related disease;
XX cardiovascular disease; brain; neuronal associated disease; metabolic;
XX inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
XX antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;
XX Nkak2.
XX Homo sapiens.
OS 72.
XX MO2004006838-A2.
XX 22-JAN-2004.
XX 15-JUL-2003; 2003MO-US021730.
XX 15-JUL-2002; 2002US-0395632P.
XX (SUGR-) SUGEN INC.
XX Whyte D, Manning G, Caenepeel S;
XX WPI; 2004-122753/12.
XX N-PSDB; ADJ96554.
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX Claim 1; SEQ ID NO 77; 366bp; English.
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytostatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
SQ Sequence 672 AA;
Query Match 96.3%; Score 1293; DB 8; Length 672;
Best Local Similarity 95.7%; Pred. No. 7.5e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETLGKGTGKVKKARSSGRVAIKSIRKDKDEODLHIREIRIEMSLNPHI 60
DB 97 YEFLETLGKGTGKVKKARSSGRVAIKSIRKDKDEODLHIREIRIEMSLNPHI 156
QY 61 IAIHEVFENSSKIVIMEVYASRGDLVDYISERPRLSERDARHFRQIVSAHYCHONGIV 120
DB 157 IAIHEVFENSSKIVIMEVYASRGDLVDYISERQOLSREARHFRQIVSAHYCHONGIV 216
QY 121 HRDLKLENTLLDANGNINKIADFGLSNLYHGKFLQTFCCSPLVYASPEIVNGKPYGPEVD 180
DB 217 HRDLKLENTLLDANGNINKIADFGLSNLYHGKFLQTFCCSPLVYASPEIVNGKPYGPEVD 276
QY 181 SWSLGVLLYLVHGTMPEFDGQDHKTIVKQISNGAYRPPKPSDACGLIRWLWVNPTRRA 240

Db 277 SMSLGVLYLVHGTMPPDGDHDKILVKQISNGAYREPPKSPDACGLIRWLMLVNPTRRA 336
QY 241 TLEDVASHMWVNM 253
Db 337 TLEDVASHMWVNM 349

RESULT 15
AAE16266
AAE16266 standard; protein; 594 AA.

AAE16266;
26-MAR-2002 (first entry)
Human kinase PKIN-12 protein.

Human kinase PKIN-12; cancer; leukaemia; adenocarcinoma; osteoporosis;
immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
congestive heart failure; ischaemic heart disease; lung tumour; gout;
fatty liver; Niemann-Pick's disease; gene therapy.

XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT Domain 19..269
FT Domain /note="Bukaryotic protein kinase domain"
FT Domain 22..260
FT Domain /label= Protein_kinase_domain
FT Domain 23..260
FT Domain /label= Protein_kinase_domain
FT Domain 24..260
FT Domain /label= Protein_kinase_domain

XX WO200196547-A2.
XX
XX 20-DEC-2001.
XX
XX 14-JUN-2001; 2001WO-US019444.
XX
XX 15-JUN-2000; 2000US-0212073P.
XX 23-JUN-2000; 2000US-0213467P.
XX 30-JUN-2000; 2000US-0215651P.
XX 07-JUL-2000; 2000US-0216605P.
XX 13-JUL-2000; 2000US-0218572P.
XX 25-AUG-2000; 2000US-0228056P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y,
PI Gandit AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR,
PI Ramkumar J, Griffin JA, Kearney L, Buford N, Nguyen DB, Tang YT,
PI Baughn MR, He A, Thornton M, Hafalla A, Patterson C, Gururajan R,
PI Lo TP, Khan F, Reclison SA, Azimzai Y, Policky JL, Ding L,
PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX
XX WPI; 2002-090207/12.
XX N-PSDB; AAD26459.
XX
XX New polypeptides, useful for diagnosing, treating or preventing disorders
XX of growth and development, cardiovascular and lipid, and diseases such as
XX cancer, comprise human kinase polypeptides.
XX
XX Claim 1, Page 152-153; 197PP; English.
XX
XX The invention relates to human kinase PKIN proteins and their

CC corresponding cDNAs. A composition containing PKIN agonist is useful for
CC treating a disease or condition associated with decreased expression of
CC PKIN and a composition comprising PKIN antagonist is useful for treating
CC a disease or condition associated with overexpression of PKIN. The
CC disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
CC hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity
CC of a test compound and in gene therapy. The present sequence is human
CC PKIN-12 protein

XX
XX
SQ Sequence 594 AA;
Query Match 95.8%; Score 1286; DB 5; Length 594;
Best Local Similarity 95.3%; Pred. No. 3.7e-132;
Matches 241; Conservative .6; Mismatches 6; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKYVKKARESSGRVVAIKIRKDKIKDEODLLHIREIIMSSLNHPH 60
Db 19 YEFLETLGKGTGKYVKKARESSGRVVAIKIRKDKIKDEODLLHIREIIMSSLNHPH 78

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARRFPQIVSALHYCHQNGIV 120
Db 79 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERARRFPQIVSAVHYCHQNRV 138

QY 121 HRDLKLENTLLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
Db 139 HRDLKLENTLLDANGNIKADPGLSNLYHQGFLOTFCGSPLYASPEIYNGKPYGPEVD 198

QY 181 SMSLGVLYLVHGTMPPDGDHDKILVKQISNGAYREPPKSPDACGLIRWLMLVNPTRRA 240
Db 199 SMSLGVLYLVHGTMPPDGDHDKILVKQISNGAYREPPKSPDACGLIRWLMLVNPTRRA 258

QY 241 TLEDVASHMWVNM 253
Db 259 TLEDVASHMWVNM 271

Search completed: May 11, 2005, 14:20:08
Job time : 67.3982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 15, 2005, 11:55:01 ; Search time 303.196 Seconds
(without alignments)
10216.077 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atgagagtcggtgagcttact.....gaatctgctcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/prodata/1/ina/5A COMB.seq:*
2: /cgm2_6/prodata/1/ina/5B COMB.seq:*
3: /cgm2_6/prodata/1/ina/6A COMB.seq:*
4: /cgm2_6/prodata/1/ina/6B COMB.seq:*
5: /cgm2_6/prodata/1/ina/6C COMB.seq:*
6: /cgm2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2902	US-09-579-664B-4	Sequence 4, Appl1
2	1893	100.0	2902	US-10-355-975A-4	Sequence 4, Appl1
3	1355.4	71.6	2501	US-09-799-451-215	Sequence 215, App
4	204.2	10.8	2224	US-09-949-016-2384	Sequence 2384, Ap
5	204.2	10.8	2224	US-09-949-016-2385	Sequence 2385, Ap
6	204.2	10.8	2950	US-09-949-016-1546	Sequence 1546, Ap
7	204.2	10.8	2950	US-09-949-016-1547	Sequence 1547, Ap
8	202.6	10.7	2175	US-09-984-890-1	Sequence 1, Appl1
9	202.6	10.7	2175	US-10-274-194-1	Sequence 1, Appl1
10	189.2	10.0	484	US-09-270-767-1257	Sequence 1257, Ap
11	189.2	10.0	484	US-09-270-767-16539	Sequence 16539, A
12	186.4	9.8	2908	US-09-930-181-1	Sequence 1, Appl1
13	182.4	9.6	2652	US-08-557-006C-39	Sequence 39, Appl
14	182.4	9.6	2761	US-08-557-006C-24	Sequence 24, Appl
15	181.2	9.6	1742	US-08-557-006C-38	Sequence 38, Appl
16	179.8	9.5	2698	US-08-677-298-1	Sequence 1, Appl1
17	179.8	9.5	2912	US-09-949-016-3363	Sequence 2163, Ap
18	178.2	9.4	2914	US-09-949-016-343	Sequence 2343, App
19	176.4	9.3	2899	US-09-774-528-112	Sequence 112, App
20	174	9.2	3364	US-09-930-181-3	Sequence 3, Appl1
21	173	9.1	2337	US-10-003-690-3	Sequence 3, Appl1
22	173	9.1	2337	US-10-116-326-1	Sequence 1, Appl1
23	173	9.1	2980	US-10-003-690-1	Sequence 1, Appl1
24	172.2	9.1	1539	US-09-633-328B-1	Sequence 1, Appl1
25	172	9.1	3609	US-09-799-875-6	Sequence 6, Appl1
26	172	9.1	5983	US-09-799-875-4	Sequence 4, Appl1
27	171.4	9.1	1647	US-09-101-146-44	Sequence 44, Appl

28	165.6	8.7	1229	US-09-359-161-4	Sequence 4, Appl1
29	163.6	8.6	1747	US-08-557-006C-44	Sequence 44, Appl
30	161.6	8.5	1863	US-09-949-016-4303	Sequence 4303, Ap
31	160.4	8.5	1507	US-09-523-849-4	Sequence 4, Appl1
32	160.4	8.5	1544	US-09-734-673-1	Sequence 1, Appl1
33	160.4	8.5	2060	US-09-523-849-1	Sequence 1, Appl1
34	158	8.3	1736	US-08-557-006C-37	Sequence 37, Appl
35	158	8.3	1783	US-08-557-006C-36	Sequence 36, Appl
36	155.6	8.2	1051	US-09-523-849-3	Sequence 3, Appl1
37	154	8.1	2112	US-10-116-326-5	Sequence 5, Appl1
38	149.4	7.9	1599	US-09-256-465-1	Sequence 1, Appl1
39	149.4	7.9	1599	US-09-167-322-3	Sequence 3, Appl1
40	149.4	7.9	1599	US-09-023-653-1004	Sequence 1004, Ap
41	147.8	7.8	1254	US-09-590-740-3	Sequence 3, Appl1
42	144.4	7.6	2181	US-09-417-197-70	Sequence 70, Appl
43	144.4	7.6	2184	US-09-417-197-138	Sequence 138, App
44	144.4	7.6	2610	US-09-212-771-1	Sequence 1, Appl1
45	144.4	7.6	2610	US-09-091-058-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1									
US-09-579-664B-4									
; Sequence 4, Application US/09579664B									
; Patent No. 6514719									
; GENERAL INFORMATION:									
; APPLICANT: Immunex Corporation									
; APPLICANT: Bird, Timothy A.									
; APPLICANT: Varca, G. Duke									
; APPLICANT: Martin, Unja									
; APPLICANT: Anderson, Dirk M.									
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES									
; FILE REFERENCE: 2923-A									
; CURRENT APPLICATION NUMBER: US/09/579,664B									
; CURRENT FILING DATE: 2000-05-26									
; NUMBER OF SEQ ID NOS: 36									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 4									
; LENGTH: 2902									
; TYPE: DNA									
; ORGANISM: Mus musculus									
US-09-579-664B-4									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGAGTCGCTGAGCTTACTCCAGCGCCGAGCGCTCCGCGCTCCGCGCTGCC	60						
DB	123	ATGAGTCGCTGAGCTTACTCCAGCGCCGAGCGCTCCGCGCTCCGCGCTGCC	182						
QY	61	TCGAGAGCGCCCGCGCGCTGCGGAGCGGCTCATGATGCTTAACCTTGATGAAG	120						
DB	183	TCGAGAGCGCCCGCGCGCTGCGGAGCGGCTCATGATGCTTAACCTTGATGAAG	242						
QY	121	AACGAGCGGCTGAAGCGGCATCAACAACAACAACCTGGGCAACGCTTCAAGTTCTTG	180						
DB	243	AACGAGCGGCTGAAGCGGCATCAACAACAACAACCTGGGCAACGCTTCAAGTTCTTG	302						
QY	181	GAGAGCGTGGGCAAGGCACTTACGGGAAAGGTGAAGAGGACGAGAGAGCTCGGGCGT	240						
DB	303	GAGAGCGTGGGCAAGGCACTTACGGGAAAGGTGAAGAGGACGAGAGAGCTCGGGCGT	362						
QY	241	CTGTGCGCCATCAAGTCCATCAGAAAGACAAATCAAAATAGACAGATCTGCTGCAC	300						
DB	363	CTGTGCGCCATCAAGTCCATCAGAAAGACAAATCAAAATAGACAGATCTGCTGCAC	422						
QY	301	ATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTCATTCAT	360						
DB	423	ATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTCATTCAT	482						

QY 241 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGATCTGCTCAC 300
DB 363 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGATCTGCTCAC 422
QY 301 ATACGAGGAGATGATGATCATGTCTTCACTCAACACCCCATCATTTGCCATCAT 360
DB 423 ATACGAGGAGATGATGATCATGTCTTCACTCAACACCCCATCATTTGCCATCAT 482
QY 361 GAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 483 GAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
QY 421 CTGATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 543 CTGATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 481 CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 603 CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
QY 541 AAGCTGAAAACATCCCTTCTAGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 663 AAGCTGAAAACATCCCTTCTAGATGATGATGATGATGATGATGATGATGATGATGAT 722
QY 601 TCCAACTGTACCAAAAGGCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTCTACGCC 660
DB 723 TCCAACTGTACCAAAAGGCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTCTACGCC 782
QY 661 TCGCCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 783 TCGCCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
QY 721 GGGGTTCTCTGTATCATCTCTGTGTGATGATGATGATGATGATGATGATGATGAT 780
DB 843 GGGGTTCTCTGTATCATCTCTGTGTGATGATGATGATGATGATGATGATGATGAT 902
QY 781 ACACTGGTGAAGCAATCATGTAACGGGGCTTACCGGTGAGCGCCCAAGCGTCCGATGCC 840
DB 903 ACACTGGTGAAGCAATCATGTAACGGGGCTTACCGGTGAGCGCCCAAGCGTCCGATGCC 962
QY 841 TGTGGCCTGATCCGGTGGCTGTTAATGATGTAACCCACCCGTCGGGCCCAACTGAGAGAT 900
DB 963 TGTGGCCTGATCCGGTGGCTGTTAATGATGTAACCCACCCGTCGGGCCCAACTGAGAGAT 1022
QY 901 GTAGCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1023 GTAGCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 961 CTGCGTGAAGGATGAGGACCCCTAGTGTGATCTTTGGCGGGCTTCATTTGGGAGATGGTTA 1020
DB 1083 CTGCGTGAAGGATGAGGACCCCTAGTGTGATCTTTGGCGGGCTTCATTTGGGAGATGGTTA 1142
QY 1021 CGTGGCTCTCTGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1143 CGTGGCTCTCTGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1202
QY 1081 CAGGTGCGGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1203 CAGGTGCGGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
QY 1141 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1263 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
QY 1201 TCTGCGCTTGGCAAGACAGCTTAACTTCCGAAGAGCATTTCAAAGAAAAGTCTCT 1260
DB 1323 TCTGCGCTTGGCAAGACAGCTTAACTTCCGAAGAGCATTTCAAAGAAAAGTCTCT 1382
QY 1261 AACTGCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1383 AACTGCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442

QY 1321 GGGCAGCCTGTCCCTGTGATATCCCTGTCCAGAGAAAGGCAATCCTTAAGAATCTCGA 1380
DB 1443 GGGCAGCCTGTCCCTGTGATATCCCTGTCCAGAGAAAGGCAATCCTTAAGAATCTCGA 1502
QY 1381 CAGGTGAATCTGATTAATCTCTCTCCAGAGCCAGAGGATGATGATGATGATGATGAT 1440
DB 1503 CAGGTGAATCTGATTAATCTCTCTCCAGAGCCAGAGGATGATGATGATGATGATGAT 1562
QY 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1563 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
QY 1501 CTCTCTCTCAACCGCAAGGCAATTTCTAAATCTAATGCAATTTCTCCGCAAGCTTGA 1560
DB 1623 CTCTCTCTCAACCGCAAGGCAATTTCTAAATCTAATGCAATTTCTCCGCAAGCTTGA 1682
QY 1561 GAAGGCACTAACCTCTGACCTTTGGCTCCCTGAGCAATGAGCTCTCCATCTCGCA 1620
DB 1683 GAAGGCACTAACCTCTGACCTTTGGCTCCCTGAGCAATGAGCTCTCCATCTCGCA 1742
QY 1621 GCCGGCCAGCCGCCCTCTGAGGGGCTGTGATGATGATGATGATGATGATGATGATGAT 1680
DB 1743 GCCGGCCAGCCGCCCTCTGAGGGGCTGTGATGATGATGATGATGATGATGATGATGAT 1802
QY 1681 TTGACCAATTTGACCTTCTGAAAGCTTTCTCCGAAACCCCACTGAGGGGCTGTGCT 1740
DB 1803 TTGACCAATTTGACCTTCTGAAAGCTTTCTCCGAAACCCCACTGAGGGGCTGTGCT 1862
QY 1741 GTGACCACTGAGGGGCTTTGAGCAAGCTCTCTGAAAGCTTTGAAAGCTTTGAAAGCT 1800
DB 1863 GTGACCACTGAGGGGCTTTGAGCAAGCTCTCTGAAAGCTTTGAAAGCTTTGAAAGCT 1922
QY 1801 GAATCCTTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1923 GAATCCTTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
QY 1861 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
DB 1983 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 2015

RESULT 3
US-09-799-451-215
Sequence 215, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, RyLe
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yuning
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Mehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Dimahe, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ. ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 215
LENGTH: 2501

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106) .. (1989)
US-09-799-451-215

```

Query Match 71.6%; Score 1355.4; DB 4; Length 2501;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

```

QY 47 CTTCCGCTGCTGCGAGAGCCGCCGCTGCGGAGCGGGCTCATCAAGTCGCTA 106
DB 140 CCACTCCCTCGCGCGAGAGCTACCCGCCGCTGGGGAGGGCTGATCAAGTCGCCA 199
QY 107 AACTCTGATGAGAGAGAGCGGCTGAAAGCGGCACATCAAAACAACCTGCGGACC 166
DB 200 AGCCCTTAATGAAAGACAGCGCGGTGAAGCGGCACACACAAACACCACTCGCGGACC 259
QY 167 GCTACGAGTTCTGAGAGACGCTGGGCAAGGCACTACGGGAGGTGAAGAGAGAGAG 226
DB 260 GCTACGAGTTCTGAGAGACCTGAGGCAAGGCACTACGGGAGGTGAAGAGAGAGAG 319
QY 227 AGAGCTGGGGCGCTCTGTGGCCATCAAGTCATCAGGAAAGACAAATCAAGATAGC 286
DB 320 AGAGCTGGGGCGCTGTGGCCATCAAGTCATCAGGAAAGACAAATCAAGATAGC 379
QY 287 AGATCTGCTGCAATACGAGGAGATTTAGATCATGTCTTCACTCAACCCCA 346
DB 380 AAGATCTGATGCAATACGAGGAGATTTAGATCATGTCTTCACTCAACCCCA 439
QY 347 TCATTCGCATCATGAAGTGTGGAATGAGACAGCAAGATTGATGATGATGATG 406
DB 440 TCATTCGCATCATGAAGTGTGGAATGAGACAGCAAGATTGATGATGATGATG 499
QY 407 CCAAGCCGAGCGATCTGTATGATTAATCAATGAGCGGCAAGCTGATGAGCGGAGC 466
DB 500 CCAAGCCGAGCGATCTGTATGATTAATCAATGAGCGGCAAGCTGATGAGCGGAGC 559
QY 467 CCAAGCGATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCAACGAGATCG 526
DB 560 CTAGGCAATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCAACGAGATCG 619
QY 527 TTACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCAATGGAAGAAATCAAGATTG 586
DB 620 TTACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCAATGGAAGAAATCAAGATTG 679
QY 587 CTGACTTTGGCTCTTCCAACTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGAG 646
DB 680 CTGACTTTGGCTCTTCCAACTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGAG 739
QY 647 GGCCTCTTACGCTGCTGAGATGATCAACGGAAGCCCTTATGTGGGCCAGAGGTGG 706
DB 740 GGCCTCTTACGCTGCTGAGATGATCAACGGAAGCCCTTATGTGGGCCAGAGGTGG 799
QY 707 ACAAGTGTGTCTGGGGGCTCTCTGTACATCTCGGTGCAATGGAACATCCCTTTAGC 766
DB 800 ACAAGTGTGTCTGGGGGCTCTCTGTACATCTCGGTGCAATGGAACATCCCTTTAGC 859
QY 767 GGCAGGATCATAAACACTGGTGAAGCAATCAGTACGAGGAGCTTACGTTAGAGCGGCCA 826
DB 860 GGCAGGATCATAAACACTGGTGAAGCAATCAGTACGAGGAGCTTACGTTAGAGCGGCCA 919
QY 827 AGCGGTCCGATGCTGTGGCTTATCCGGTGTGTTAATGTGAACCCCAACCGGTGGG 886
DB 920 AACCTCTGTATGCTGTGGCTTATCCGGTGTGTTAATGTGAACCCCAACCGGTGGG 979
QY 887 CCAACCTGAGGATGATGAGCAATGATGAGGAGCTTACGTTAGAGCGGAGTGC 946
DB 980 CCAACCTGAGGATGATGAGCAATGATGAGGAGCTTACGTTAGAGCGGAGTGC 1039
QY 947 GGAAGAGAGAGCGCTGTGAGGAGTGGGACCTTATGAGTACTTTGGCGCGGCTCCA 1006

```

```

DB 1040 GAGAGCAGAGAGGCTCCGATGAGGAGTGGGACCCCTGGCAATGACTGTGCCGGCTCCA 1099
QY 1007 TGGGAGCTGTGTTACGTGCTCTCGGCGCCCTCTCTGAGAAATGAGAGCAAGTGTGCA 1066
DB 1100 TGGCTGAGCTGCGCGCGCTTCTCTCGCGCCCTCTCTGAGAAATGAGGAGTGTGCA 1159
QY 1067 GCTTCTTCAAGCAGACGAGCGGAGTGGGAGCACTGATCACTGGGCTGAGCGGAGC 1126
DB 1160 GCTTCTTCAAGCAGACGAGCTGTGTGGGAGAGCAACCCCTGGCTGAGCGGAGC 1219
QY 1127 ATTCTTTAAGAGTCCCGAAGAGAAATGACATGAGCTCAAAATTTGCAAGTACCCGG 1186
DB 1220 ATTCTTTAAGAGTCCCGAAGAGAAATGACATGAGCTCAAAATTTGCAAGTACCCGG 1279
QY 1187 CTGAGATACCTTTCTTCTGCTGCGGAGAGCAAGCTTAAAGCTTCCGAAAGCAATTC 1246
DB 1280 CTGATGACATGCGCCATTCGCTGCGGAGAGCAAGCTTAAAGCTTCCGAAAGCAATTC 1339
QY 1247 AGAAAGATCTTACCTGCTGAGGAGAGTACAGAGAGCCCTCAAGAACTGAGACCGG 1306
DB 1340 AGAAAGATCTTACCTGCTGAGGAGAGTACAGAGAGCCCTCAAGAACTGAGACCGG 1399
QY 1307 TGCCTGATACCTCCAGGAGAGCTGTGCTGCTGATTCCTGCTCCCAAGAGAGAGATCC 1366
DB 1400 TGCCTGAGAGCCCAAGGAGAGGCTGCGC-----CCCTGCTCCCAAGAGAGAGATTC 1450
QY 1367 TTAAGAAATTCACAGAGCTGATCTGTGTTACTGATCTCTTCCAGAGCCAGAGTCTG 1426
DB 1451 TCAAGAAAGCCCAAGAGAGCTGTGCTGATCTCTTCCAGAGCCAGAGTCTG 1510
QY 1427 GGAAGCTTTAAGAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1486
DB 1511 GGAAGCTTTAAGAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1570
QY 1487 CACAGGCTTCAAGGCTCTCTTCCAGAGAGGCAATTCATCAAAATGAGCAAGTTCT 1546
DB 1571 CGAAGCTTCAAGGCTCTCTTCCAGAGAGGCAATTCATCAAAATGAGCAAGTTCT 1630
QY 1547 CCGGCAAGCTTTAAGAGAGCAATCCCTTACAGACTTTGGCTCTTCCAGAGCAATGGCT 1606
DB 1631 CCGGCAAGCTTTAAGAGAGCAATCCCTTACAGACTTTGGCTCTTCCAGAGCAATGGCT 1690
QY 1607 CTTCCCATCTGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1666
DB 1691 CACTGCGCCCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1750
QY 1667 TGTCTCCGAGTCTTGTGACCAATGAGACTTGTGAGAGTCTTCCGAAACCCCACTGA 1726
DB 1751 TGTCTCCGAGTCTTGTGACCAATGAGACTTGTGAGAGTCTTCCGAAACCCCACTGC 1810
QY 1727 GGGGCTGTGTGTGTGAGCAACCTGAGGAGGCTTGAAGAGCTCTCCCTCAGAG----- 1780
DB 1811 GGGGCTGTGTGTGTGAGCAACCTGAGGAGGCTTGAAGAGGCTCTCCCTCAGAG----- 1870
QY 1781 -----GTCTGAAGGCAATGTGAGAGAGATCTTGTGGGAGATGAGTCTTTCTTGAAG 1834
DB 1871 GAAAGCTGCTGAGAGGCTGAGGAGAGATCTTGTGGGAGATGAGTCTTTCTTGAAG 1930
QY 1835 ACTGCAAGAGGAGTGTGAGCTGAGCTTACAGCAAGCCCTTAAAGATCTCTCAAGAGTCAAGC 1893
DB 1931 ACTGCAAGAGGAGTGTGAGCTGAGCTTACAGCAAGCCCTTAAAGATCTCTCAAGAGTCAAGC 1989

```

```

RESULT 4
US-09-949-016-2384
; Sequence 2384, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

US-09-949-016-2385
; Sequence 2385, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/223,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2384

Query Match 10.8%; Score 204.2; DB 4; Length 2224;
Best Local Similarity 55.0%; Pred. No. 1.2e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGGAGACGCTGGGCAAGGCACTACCGGAGGTGAAGAGCAGAGA 227
DB 57 CTACCGGCTCTCAAGACATTTGGCAAGGTAATTTTCCAGGTGAAGTTGGCCGACA 116
QY 228 GAGCTCGGGGCTGTGGCCATCAAGTCCATCAAGAAACAATCAAAATGAGCA 287
DB 117 CATCTGACGGAAGAGGTAGCTGTGAAGATCAATGACAACTCAACTGAACTCTC 176
QY 288 GGATCTGTCGACATACGAGAGGAGATTGAGATCATGCTTCACTCAACCCCAAT 347
DB 177 CAGCTTCAGAAATATTTCCGGAAGTAAGTAATGAAGTTTGAATCATCCCAAT 236
QY 348 CATTGCATCATGAAGTGTGAAGTAAGCAGCAAGATTTGATGTGATGAGATTC 407
DB 237 AGTTAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 296
QY 408 CAGCGAGGCGATCTGTATGATTAATCATGAGCGGCAAGGCTGAGTGAAGCGG 467
DB 297 TAGTGGGAGAGGATTTGATTAATCATGAGGCTCATGAGGATGAAGAAAGAGGC 356
QY 468 CAGCATTTCTTCCGACAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
DB 357 TCGAGCCAAATTCGCGCAGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416
QY 528 TCACGAGATCTCAAGCTGAGAAACATCTTCTGATGTCGATGAGAAACATCAAGTTC 587
DB 417 CCATAGAGACTTAAGGAGCAAAACCTGCTTGTGATGCTGATTAAGCAATCAAGTTC 476
QY 588 TGACTTTGGCTCTCAACCTGTAACAACAAAGCAAGTTCTCAAGCGTTCTGTGGAG 647
DB 477 AGACTTTGGCTTCAAGCAATGATTAATCACTTTGGGAAACAGCTGACACTTCTGTGGAG 536
QY 648 CCGCTCTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
DB 537 TCCCCCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
QY 708 CAGCTGTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
DB 597 TGTGTGAGGCTTGAAGATTAATCTCTATACCTGCTGAGCGGATCCCTTTTGAAG 656
QY 768 GCAGAGATCAATTAACCTGTAAGCAAAATCAGTAACGCGGCTTAACGTCGAGCCCA 827
DB 657 ACAGAACTCAAGAGAGCTGCGGAAACGCGTACTGAGGGGAAATATATGAGACCGGAGTGA 716
QY 828 GCCGTCCGATGCTGTG---GCTGATCCGCTGTTAATGTTGAACCCCAACCGCTG 884
DB 717 CATGTCAACGAGCTGTGAAGCTGCTTAAGAAATTTCTCATTTCTTAATCCAGAGAG 776
QY 885 GGCACACATGAGATGATGAGCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 933
DB 777 AGGACTTTAGAGCAATATGAAAGATGATGATGATGATGATGATGATGATGATGAT 825

RESULT 5

US-09-949-016-2385
; Sequence 2385, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/223,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2385

Query Match 10.8%; Score 204.2; DB 4; Length 2224;
Best Local Similarity 55.0%; Pred. No. 1.2e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGGAGACGCTGGGCAAGGCACTACCGGAGGTGAAGAGCAGAGA 227
DB 57 CTACCGGCTCTCAAGACATTTGGCAAGGTAATTTTCCAGGTGAAGTTGGCCGACA 116
QY 228 GAGCTCGGGGCTGTGGCCATCAAGTCCATCAAGAAACAATCAAAATGAGCA 287
DB 117 CATCTGACGGAAGAGGTAGCTGTGAAGATCAATGACAACTCAACTGAACTCTC 176
QY 288 GGATCTGTCGACATACGAGAGGAGATTGAGATCATGCTTCACTCAACCCCAAT 347
DB 177 CAGCTTCAGAAATATTTCCGGAAGTAAGTAATGAAGTTTGAATCATCCCAAT 236
QY 348 CATTGCATCATGAAGTGTGAAGTAAGCAGCAAGATTTGATGTGATGAGATTC 407
DB 237 AGTTAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 296
QY 408 CAGCGAGGCGATCTGTATGATTAATCATGAGCGGCAAGGCTGAGTGAAGCGG 467
DB 297 TAGTGGGAGAGGATTTGATTAATCATGAGGCTCATGAGGATGAAGAAAGAGGC 356
QY 468 CAGCATTTCTTCCGACAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
DB 357 TCGAGCCAAATTCGCGCAGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416
QY 528 TCACGAGATCTCAAGCTGAGAAACATCTTCTGATGTCGATGAGAAACATCAAGTTC 587
DB 417 CCATAGAGACTTAAGGAGCAAAACCTGCTTGTGATGCTGATTAAGCAATCAAGTTC 476
QY 588 TGACTTTGGCTCTCAACCTGTAACAACAAAGCAAGTTCTCAAGCGTTCTGTGGAG 647
DB 477 AGACTTTGGCTTCAAGCAATGATTAATCACTTTGGGAAACAGCTGACACTTCTGTGGAG 536
QY 708 CAGCTGTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
DB 597 TGTGTGAGGCTTGAAGATTAATCTCTATACCTGCTGAGCGGATCCCTTTTGAAG 656
QY 768 GCAGAGATCAATTAACCTGTAAGCAAAATCAGTAACGCGGCTTAACGTCGAGCCCA 827
DB 657 ACAGAACTCAAGAGAGCTGCGGAAACGCGTACTGAGGGGAAATATATGAGACCGGAGTGA 716
QY 828 GCCGTCCGATGCTGTG---GCTGATCCGCTGTTAATGTTGAACCCCAACCGCTG 884

Db 717 CATGTCCACGAGCTGTGTAACCTGCTTAAGAAATTTCTATTCTTAATCCCAAGAG 776
QY 885 GGCCACACTGAGAGATGACCACTCATGTGGGTCAACTGGGTTAC 933
Db 777 AGGCACCTTAGAGCAATCATGAAGAATCGATGATGATGGGTAC 825

RESULT 6

US-09-949-016-1546
; Sequence 1546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1546
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1546

Query Match 10.8%; Score 204.2; DB 4; Length 2950;
Best Local Similarity 55.0%; Pred. No. 1.4e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGAGACGCTGCGCAAGGCACTCAAGGAAAGTTGAAGACGAGA 227
Db 464 CTACCGGCTCTCAAGACCAATGGCAAGGTAATTTGGCAAGGTGAAGTTGGCCGACA 523
QY 228 GAGCTCGGGCGCTGTGGCCATCAATCCATCAGAAAGCAAAATCAAGATGAGA 287
Db 524 CATCTGACCTGAGAAAGAGTAGCTGTGAAGATCATTTGACAAAGCTCAACTGACTCTC 583
QY 288 GATCTGCTGCACATACGAGAGGAGATGATCATCTTCACTCAACCACTCCCAT 347
Db 584 CAGCTTCAGAAACTATTCGCGAAGTAAGAAATTAAGGTTTGAATCATCCACAT 643
QY 348 CATTCGATCATGAAGTGTGTAAGATAGACGAAAGATTTGTCTATGAGATGTC 407
Db 644 AGTTAAATTTTGAAGTGAATGAGACTGAGAAACCTCTCACTTCTCATGAGTACG 703
QY 408 CAGCCGAGGAGATCTGATGATTACATCAGTGAAGCGGCAAGCTGAGTGAAGCGG 467
Db 704 TAGTGGCGAGAGATTTTATTAATCTAGTGTCTCATGAGAGATGAAGAAAAAGGCG 763
QY 468 CAGGCAATTTCTCGACAGATGCTGTGCTGCTGCACTACTGCGCAGCAGAACGGATCGT 527
Db 764 TCAGGCCAAATTCGCGAGATAGTGTCTGCTGTCAGATCATCTGACAGAAAGTTATGT 823
QY 528 TCACCGAGATCTCAAGTGAAGAAATCTTTTATGATGCCAATGAGAAATCAAGATTGC 587
Db 824 CCAATAGAGACTTAAGGCAAAAACTGCTTTGAGATGATGATGAAATCAAGATTGC 883
QY 588 TGACTTTGGCTCTCCAACTGTACCAAAAGCAAGTTCTCTCAGAGCTTTGTGGAG 647
Db 884 AGACTTTGGCTCTCAAGCATGAATTCACCTTTGGAGAACAGTGAACCTTCTGTGGAG 943
QY 648 CCTCTCTACGCTCTGCTGATGATGATCAACGAGAACCTTATGTGGGCCAGAGTGA 707
Db 944 TCCCCCTTATGTGCGCCAGAACTCTTCCAGGGGAAAAAATATGATGAGACCCGAGGTGA 1003

QY 708 CAGCTGCTCTGGGGCTTCTCTGTACATCTGTGTCAGACCAATGCCCTTTGACG 767
Db 1004 TGTGTGAGCTTAGAGATTATCTTATACATCTGTGACGGATCTGCTTTGTAGG 1063
QY 768 GCAGATCATATAACAATCTGTGAAGCAATCACTTAACGGGGCTTACCGTAGCGCCAA 827
Db 1064 ACAGAACCTTAAGAGGTGGGGAGACGGGTACTGAGGGGAAATACCGTATTCATTTCTA 1123
QY 828 GCCGTCCGATCCTGTG--GCCGTATCCGTGGCTGTTAATGTGAACCCACCGCTG 884
Db 1124 CATGTCCACGAGCTGTGAATAAAGCTGTTAAGAAATTTCTATTTATCCCAAGAG 1183
QY 885 GGCCACACTGAGAGATGACCACTCATGTGGGTCAACTGGGTTAC 933
Db 1184 AGGCACCTTAGAGCAATCATGAAGAATCGATGATGATGGGTAC 1232

RESULT 7

US-09-949-016-1547
; Sequence 1547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1547
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1547

Query Match 10.8%; Score 204.2; DB 4; Length 2950;
Best Local Similarity 55.0%; Pred. No. 1.4e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGAGACGCTGCGCAAGGCACTCAAGGAAAGTTGAAGACGAGA 227
Db 464 CTACCGGCTCTCAAGACCAATGGCAAGGTAATTTGGCAAGGTGAAGTTGGCCGACA 523
QY 228 GAGCTCGGGCGCTGTGGCCATCAATCCATCAGAAAGCAAAATCAAGATGAGA 287
Db 524 CATCTGACCTGAGAAAGAGTAGCTGTGAAGATCATTTGACAAAGCTCAACTGACTCTC 583
QY 288 GATCTGCTGCACATACGAGAGGAGATGATCATCTTCACTCAACCACTCCCAT 347
Db 584 CAGCTTCAGAAACTATTCGCGAAGTAAGAAATTAAGGTTTGAATCATCCACAT 643
QY 348 CATTCGATCATGAAGTGTGTAAGATAGCAGAGATTTGATGTCATGAGATGTC 407
Db 644 AGTTAAATTTTGAAGTGAATGAGACTGAGAAACGCTTACCTTGTCTATGAGATGACG 703
QY 408 CAGCCGAGGAGATCTGATGATTACATCAGTGAAGCGGCAAGCTGAGTGAAGCGGAGCG 467
Db 704 TAGTGGCGAGAGATTTTATTAATCTAGTGTCTATGAGAGATGAAGAAAAAGAGCG 763
QY 468 CAGGCAATTTCTCGACAGATGCTGTGCTGCTGCACTACTGCGACCAAGACGGATCGT 527
Db 764 TCAGGCCAAATTCGCGAGATAGTGTCTGCTGTCAGTACTGTCACAGAAAGTTATTTGT 823
QY 528 TCACCGAGATCTCAAGTGAAGAAATCTTTTATGATGCCAATGAGAAATCAAGATTGC 587
Db 824 CCAATAGAGACTTAAGGCAAAAACTGCTTTGAGATGATGATGAAATCAAGATTGC 883

QY 588 TGACTTGGGCTCTCCAACTGTACCAAGGCAAGTTCCTCAGAGCTTTGTGGAG 647
DB 884 AGACTTGGCTTACGACATGAATTCACCTTTGGGAAACAAGTGGACACCTTGTGGAG 943
QY 648 CCTCTCTAGAGCTTGGCTTGAATGACAAAGGAAACCTTATGTGGGCCCCAGAGTGA 707
DB 944 TCCCTCTTATGTGCGCCCAACCTTTCAGGGGCAAAAATATGATGACCCGAGGTGA 1003
QY 708 CAGCTGGTCTGTGGGCTTCTCTGTATCATCTGTGTCATGTGCAACATGCGCTTTGAAG 767
DB 1004 TGTGTGAGGCTTGAAGATTAATCTTATACATGTGTCAAGGAGATCCCTGCTTTGATGG 1063
QY 768 GCAAGATCATTAACACATGTGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGGCCAA 827
DB 1064 ACAGAACTCATAGAGCTGCGGGAAACGGTATCTGAGGGGAAATATACGTAATTCATTCTA 1123
QY 828 GCGGTCCGATGCTGTG--GCTGATCCGGTGGCTTTATGTGAACCCCAACCGCTG 884
DB 1124 CATGTCCACGAGCTGTGAACCTGTCTTAAGAAATTTCTCATTTCTTATCCCAAGAG 1183
QY 885 GGCACACATGAGATGTAGCCAGTCAATGTGGGTCAACTGGGGTTAC 933
DB 1184 AGGCACTTTAGAGCAATCATGAAGATCAGATGATGTGGGTAC 1232

RESULT 8

US-09-984-890-1
; Sequence 1, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-890-1

Query Match 10.7%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 3,4e-47;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 168 CTACGAGTTCTGTGAGACGCTGGGCAAGGGCACTAGGAGAGGTGAAGAGCAGAGA 227
DB 156 CTACCGGCTCTCAAGCAATGTGGCAAGGTAATTTGCCAAGGTGAAGTTGGCCGACA 215
QY 228 GAGCTGGGGGCTGTGGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGACA 287
DB 216 CATCTGACTGAGGAAGAGGTAGCTGTGAATCATTTGACAAAGTCAATGAACTCTC 275
QY 288 GAGTCTGTGCAATACGAGGAGATTTGATGATCTTCACTCAACACCCCAAT 347
DB 276 CAGCTTCAGAAACTATTCGCGAAGTAAGATTAATAGATTGTAATATCCCAAT 335
QY 348 CATTCGATCCATGAAGTGTGAGAAATAGCAGAAATTTGTATGTCTCATGAGATATG 407
DB 336 AGTTAAATTTATGAAATGATGAGTGAAGAAACCTTCACTGTCTATGAGATGAGC 395
QY 408 CAGCCGAGGAGATCTGTATGATTAATCATCACTGAGCGGCCACGGCTGATGAGCGGAGC 467
DB 396 TAGTGGGAGAGGATATTTGATTAATTAAGTGTCTATGAGGAGATGAAAGAAAGAGC 455
QY 468 CAGGCAATTTTCTGAGAGATCTGTGTCTGCTGTGCACTTATGCGACCAAGAGGATGCT 527
DB 456 TCAGGCCAAATTCGCGAGGATGATGTCTGTGTCAGATGATCTATGTCACCAAGAGTTATGT 515

QY 528 TCACGAGATCTCAAGCTGAAAAATCTTTTATGATGCCAATGAGAAACATCAAGATTGC 587
DB 516 CCAATAGAGACTTAAGGCAAGAAACCTGCTCTTGATGTGATATGAACATCAAGATTGC 575
QY 588 TGACTTGGGCTCTCCAACTGTACCAAGGCAAGTTCCTCAGAGCTTTGTGGAG 647
DB 576 AGACTTGGCTTACGACATGAATTCACCTTTGGGAAACAAGTGGACACCTTGTGGAG 635
QY 648 CCTCTCTAGAGCTTGGCTTGAATGACAAAGGAAACCTTATGTGGGCCCCAGAGTGA 707
DB 636 TCCCTCTTATGTGCGCCCAACCTTTCAGGGGCAAAAATATGATGACCCGAGGTGA 695
QY 708 CAGCTGGTCTGTGGGCTTCTCTGTATCATCTGTGTCATGTGCAACATGCGCTTTGAAG 767
DB 696 TGTGTGAGGCTTGAAGATTAATCTTATACATGTGTCAAGGAGATCCCTGCTTTGATGG 755
QY 768 GCAAGATCATTAACACATGTGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGGCCAA 827
DB 756 ACAGAACTCATAGAGCTGCGGGAAACGGTATCTGAGGGGAAATATACGTAATTCATTCTA 815
QY 828 GCGGTCCGATGCTGTG--GCTGATCCGGTGGCTTTATGTGAACCCCAACCGCTG 884
DB 816 CATGTCCACGAGCTGTGAACCTGTCTTAAGAAATTTCTCATTTCTTATCCCAAGAG 875
QY 885 GGCACACATGAGATGTAGCCAGTCAATGTGGGTCAACTGGGGTTAC 933
DB 876 AGGCACTTTAGAGCAATCATGAAGATCAGATGATGTGGGTAC 924

RESULT 9

US-10-274-194-1
; Sequence 1, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-274-194-1

Query Match 10.7%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 3,4e-47;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 168 CTACGAGTTCTGTGAGACGCTGGGCAAGGGCACTACGAGAGGTGAAGAGCAGAGA 227
DB 156 CTACCGGCTCTCAAGCAATGTGGCAAGGTAATTTGCCAAGGTGAAGTTGGCCGACA 215
QY 228 GAGCTGGGGGCTGTGGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGACA 287
DB 216 CATCTGACTGAGGAAGAGGTAGCTGTGAATCATTTGACAAAGTCAATGAACTCTC 275
QY 288 GAGTCTGTGCAATACGAGGAGATTTGATGATCTTCACTCAACACCCCAAT 347
DB 276 CAGCTTCAGAAACTATTCGCGAAGTAAGATTAATAGATTGTAATATCCCAAT 335
QY 348 CATTCGATCCATGAAGTGTGAGAAATAGCAGAAATTTGTATGTCTCATGAGATATG 407
DB 336 AGTTAAATTTATGAAATGATGAGTGAAGAAACCTTCACTGTCTATGAGATGAGC 395
QY 408 CAGCCGAGGAGATCTGTATGATTAATCATCACTGAGCGGCCACGGCTGATGAGCGGAGC 467
DB 396 TAGTGGGAGAGGATATTTGATTAATTAAGTGTCTATGAGGAGATGAAAGAAAGAGC 455


```

; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
; US-09-930-181-1

```

Query Match	9.8%	Score 186.4	DB 3	Length 2908
Best Local Similarity	53.5%	Pred. No. 1.6e-42		
Matches 413, Conservative	0	Mismatches 356	Indels 3	Gaps 1

QY	166	CGCTACGAGTTCTCTGGAGAGCGCTGGGCGAAGGCGACCTACCGGAAGGGAAGGACGACGA	225
Db	157	CCCTACCGCGCTGGAGAAAGACGCTGGGCGAAGGGGCAACAGAGTCTGGTGAAGCTGGGGTT	216
QY	226	GAGAGCTCGGGGCGTCTGGTGCCATCAAGTCATCAGGAAAGACAAATCAAAATGAG	285
Db	217	CACCTCGGTCACTGCGCAGAAAGGGGCGCATCAAGATGTCACCGTGGAAGTCAAGCGAG	276
QY	286	CAGAGTCTGCTCCACATACGAGAGGAGATTGAGATCATGCTTCCATCTCAACCCAC	345
Db	277	TCGGTGTGATTAAGGTGAGCGGGAGATCGCATCTGAAGCTCATTTGACACCCCGAC	338
QY	346	ATCATTTGCCATCCATGAAGTGTGGAAATAGCAGCAAGATTGTGATTTCATGAGATAT	405
Db	337	GTCTCAAAAGCTCAGACGACTTTATGAAACAAAATTTTGAACCTGGTGTCTAGAACAC	396
QY	406	GCCAGCCGAGCGGATCTGTATGATTAATCATCAGTAGCGGCCACCGCTGAGTAGCGGGAC	465
Db	397	GTGTCAAGGTGGAGCTCTTCACATCTCGTGTAAAGAGGGAGGGCTGACGCTTAAGGAG	456
QY	466	GCCAGGCAATTTCTTCCGACAGATCGATCGTGGCCCTGCACTACACGACACGGAACGGGATC	525
Db	457	GCTCGGAAGTTCTTCCGAGATCATCTCTGCGCTGACCTTCTGACACGCCACTCCATA	516
QY	526	GTTCAACCGAGATCTCAAGCTGAAAAATCTCTTGAATGCAATGGAACATCAAGATT	585
Db	517	TGCCACAGGAGATCTGAACCTGAAAACTCTGCTGAGACGAAAGAACATATCCGATC	576
QY	586	GCTGACTTTGGGCTCTCCAACTGTACACAAAGGCAAGTTTCTCCAGACGTTCTGTGGG	645
Db	577	GCAAGCTTTGGGATGCGCTGCTCGAGGTTGGGACAGCTGTTGGAGACCACTGTGGG	638
QY	646	AGCCCTCTCTACGCGCTCGCCTGAGATGTCACAGGGAAGCGCTATGNGGGGCCAGAGGTG	705
Db	637	TCCCCCACTAAGCTCTGCCCCAGAGTGATTCGGGGGGGAAGATATGACGCGCGGAAGCG	696
QY	706	GACAGCTGTCTCTGGGGCTTCTCTGTACATCTGTGTGATGACGACCAATGCGCTTTGAC	765
Db	697	GACGTGTGAGCTGCGGGCGCTCATCTGTTTCGCTTGTGTGGGGGCTGCGCTTGCAC	756
QY	766	GGGCGAGGATCAATAAACACTGGTGAAGCAAAATCAGTAAACGGGGCTTAACCGTAGCGCGCC	825
Db	757	GATGACAACTTGCGACAGCTGCTGGAGAAAGGTAAAGGGGGCGGTTCACATATCGGCAC	818
QY	826	--AAAGCGTCCGATGCCCTGTGGCTCGATTCGGTGGCTGTTAATGTGTGAACCCACCCGT	882
Db	817	TTTATCCCGCCGACTGCGCAGAGTGTGTATCGGGGATGATTCAGAGGTGAAGCGCGACGCG	876
QY	883	CGGGCGACACTGGAAGATGTAGCAATCATTTGTGGTCAACTGGGGTTACA	934

Db 877 CGCCTCAGGCTAGAGCACATT CAGAAACACATATGTTATAGGGGCAAGA 928

RESULT 13

US-08-557-006C-39
; Sequence 39, Application US/08557006C

; GENERAL INFORMATION:

APPLICANT: Beri, Rajindar K.

APPLICANT: Forder, Robert A

TITLE OF INVENTION: NUCLEIC

FILE REFERENCE: NGAP/PHM37588/UST

CURRENT APPLICATION NUMBER: US/08/

CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: BCT/GB94/01083

PRIOR FILING DATE: 1994-05-20

PRIOR APPLICATION NUMBER: GB 9310489.1

PRIOR FILING DATE: 1993-05-21 ;

PRIOR APPLICATION NUMBER: GB 9318010.7

NUMBER OF SEQ TD NOS: 44

SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 39

; LENGTH: 2652

TYPE: DNA

```

; ORGANISM: KAL
:
: FEATURE:

```

NAME/KEY: gene

LOCATION: (1) .. (1747)

OTHER INFORMATION: Full letter

OTHER INFORMATION: Fragment

OTHER INFORMATION: NUCLEONIC
US-08-557-006C-39

Query Match 9.6%; Score 182.4; DB 3; Length 2652;

Best Local Simi

Matches 437; conservative 0; mismatches 381; indels 6; gaps 2

OY	120	GAACGAGCGCGGTGAAAGCGGCACTATCAACAAACAACCTGGCGGACCGCTCTAGATTCT	179
Db	4	GAACTGGCTTGAAAGCAGAAAGCAAGACGGCGTGTGAAGATTCGGAACACTACGTGCTGGG	63
OY	180	GGAGACGCTGGGGCAAGGGCACTTACGGGAAGTGAAGAGCAAGAGAGCTCG--GG	236
Db	64	GGACACCTCTGGGCGTCGGCACTTCGGCAAAAGTGAAGATTGAGAAACATCAATTGACAGG	123
OY	237	GGGTCTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAAATGAAGTGAAGCAGATCTGCT	296
Db	124	CCATTAAGTGGCAGCTTAAAGATCTTAAATATGACAGAAAGATTGCGAGTTTGAATGTTGTTGG	183
OY	297	GCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCACTCATATTGCCAT	356
Db	184	AAAAATTAAACGAAATTTCAAATCTTAAACTCTTCGTATCTCTCATATTATCAAACT	243
OY	357	CCATGAAGTGTGGAATATGACAGCAAGATTGTGATTTGATTCATGAGATATGCGACCGAGG	416
Db	244	CTACCAAGTATCAGCACTCCAAACAGACTTTTATGTTAATGGAATATGTCTGAGAG	303
OY	417	CGATCTGATATTAATTCATCAATGAGCGGCAACGGCTGAGTGAAGCGGGAACGCCAGCATTT	476
Db	304	TGAATTTGTCACATCAATCTGTAAACACGGGAGGTTGMAAGGTGAAGCTCGCGGCT	363
OY	477	CTTCGACAGATCGTGTCTGGCCCTGCACCTACGTGCAACAGAAACGGGATCGTTCCACGAGA	536
Db	364	CTTCGACGAGATTTCTGTCTGCGCGTGACATCTGTCAACGACACATGTTGTTCACAGGGA	423
OY	537	TCTCAAGCTGAAAAACATCCTTCTTGAATGCCAATGAAACATCAAGATTGTCTGACTTTGG	596
Db	424	CCTGAAGCAAGAACGTGTGCTGGAGCGCCCAATGAATCTTAAAGATATGACTGACTTGG	483
OY	597	CCTCTCAACCTGTATCCACAAGGCAGTTCTCTCAAGCTTCTGTGTGGAGCCCTCTCTTA	656

Db 484 ACTCTAATATATGATGTCAGATGCTGAATTTCTACGAAGTACTGTGTGATGCCAAATTA 543
Qy 657 CGCCTCCGCTGAGATAGTCAACGGAAAGCCCTATATGTGGGCCCAAGAGTGAACAGCTGCTC 716
Db 544 TGCAGCACCGGAGTCACTTCAGGAAGGCTGTATGCGGGTCTTGAAGTATCTGGAG 603
Qy 717 TCTGGGGGTTCTCTGTACATCCGTGTGATGAGCAACATGCGCTTTTGAACGGGCAAGATCA 776
Db 604 CTGTGGGTATCTGTATGCTGTATGCTTCTGTGGCAACCTCCGTTTCAAGATGAGCAGT 663
Qy 777 TAAACACTGTGAGCAAAATCAGTAAACGGGCTTACCGTGAAGCGCCCAAGCCGTGCA 836
Db 664 GCTTACGCTCTTTAAGAAAGATCCGAGGGGCTGTGTCTTACATCCCGAGATATCTCAACG 723
Qy 837 TGCCTGTGGC--CTGATCCGGTGGCTGTATATGTGTGAACCCCAACCGTGGGCAACT 893
Db 724 TTTATATGCCACTCTGTGTGATGACATGCTGCAAGTGGACCCCTTGAAGCGAGCACTAT 783
Qy 894 GGAGGATGATGCCATGTGTGGGTCAACTGGGGTTACACCA 937
Db 784 CAAGACATACGAGAGCATGATGTTTAAACAGGATTTGCCA 827

RESULT 14

US-08-557-006C-24
Sequence 24, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Bertl, Rajindar K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PM37588/UST
CURRENT APPLICATION NUMBER: US/08/557, 006C
PRIOR FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR FILING DATE: 1993-05-21
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 2761
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat
US-08-557-006C-24

Query Match 9.6%; Score 182.4; DB 3; Length 2761;
Best Local Similarity 53.0%; Pred. No. 2.2e-41;

Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;

Qy 120 GAAGCAGCGGTGAAGGGGCAACATCAACAACCTGGGCGCTACGAGTTCT 179
Db 20 GAACATGGCTGAGAAAGCAGACGACGGCGTGTGAAGATGGACACTACGTCGGG 79
Qy 180 GGAGACCTGGGCAAGGGCACTTACGGGAAGTGAAGAAGCAGAGAGCTG---GG 236
Db 80 GGACACCTGGGCGTGGCACTTTCGGCAAGTGAAGATGGAAACATCAATTGACAGG 139
Qy 237 GCGTCTGGTGGCCATCAATGTCATCAGAAAGACAAATCAAAAGATGAGCAGATGCT 296
Db 140 CCATTAAGTGGCAGTTAAGATCTTAATAGACAGAAATTCGCACTTTAGATGTTGG 199
Qy 297 GCAATATCGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCAT 356
Db 200 AAAAATTAACGAGAAATTTCAAAATCTTAAACTCTTTCGTCATCTCATATTAACAAC 259
Qy 357 CCATGAAGTGTGTAAGATAGCAGCAAGATTTGATTTGTCTATGAGATGCGACCGGAG 416

Db 260 CTACCAAGTATACGACTCTCAACAGACTTTTATATGATGATGATATGTCTGAGG 319
Qy 417 CGATCTGTATGATTAATACATGATGAGCGGCCAGCTGATGAGCGGCCAGGCACTTT 476
Db 320 TGAATGTGTGACACTACTGTAAACACGGGAGGTTGAAGAGTGAAGTGAACCTGCCGCT 379
Qy 477 CTTCGACAGATGCTGTCTGCTCCCTGACATCTGCGACACAGAACGGGATGCTTACCGAGA 536
Db 380 CTTCGACAGATCTGTCTGCTCCGTTGACTCTGTACAGGACACATGCTTGTCCACAGGA 439
Qy 537 TCTCAAGCTGAGAAACCTCTTCTAGATGTCANATGAAATCAATCAAGATGCTGAG 596
Db 440 CTTGAAGCCAGAGAACCTGTGTGACGCGCCAGATGAATGCTPAAGTATCTGACTTCGG 499
Qy 597 CTTCTCAACCTGTACCAACAAAGCAAGTTCTTCCAGACGTTCTGTGGAGCCCTCTCTA 656
Db 500 ACTCTTAATATATGATGATGATGATGATGATTTCTTACGAACCTAGCTGTGATTCGCAATTA 559
Qy 657 CGCCTCGCTGAGATGATCAACGGGAGCCCTTATGTGGGCCCAAGAGTGAACAGCTGTC 716
Db 560 TGCAGCACCGGAGGTCACTCAGGAAAGCTGTATGCGGGTCTGAGGTTGATATCTGAG 619
Qy 717 TCTGGGGTCTCTGTACATCTGTGTGATGATGACATGACATGCTTTGAACGGGCAAGATCA 776
Db 620 CTGTGTGTATCTCTGTATGCCCTTCTGTGTGACACCTCCGTTGACATGAGCAGCT 679
Qy 777 TAAACACTGTGAGCAAAATCAGTAAACGGGCTTACCGTGAAGCGCCCAAGCCGTGCA 836
Db 680 GCTTACGCTCTTTAAGAAAGATCCGAGGGGCTGTGTCTTACATCCCGAGATATCTCAACG 739
Qy 837 TGCCTGTGGC--CTGATCCGGTGGCTGTATATGTGTGAACCCCAACCGTGGGCAACT 893
Db 740 TTTATATGCCACTCTGTGTGATGACATGCTGCAAGTGGACCCCTTGAAGCAGCACTAT 799
Qy 894 GGAGGATGATGCCATGTGTGGGTCAACTGGGGTTACACCA 937
Db 800 CAAGACATACGAGAGCATGATGTTTAAACAGGATTTGCCA 843

RESULT 15

US-08-557-006C-38
Sequence 38, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Bertl, Rajindar K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PM37588/UST
CURRENT APPLICATION NUMBER: US/08/557, 006C
PRIOR FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR FILING DATE: 1993-05-21
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 38
LENGTH: 1742
TYPE: DNA
ORGANISM: Human AMP protein kinase
US-08-557-006C-38

Query Match 9.6%; Score 181.2; DB 3; Length 1742;
Best Local Similarity 53.1%; Pred. No. 3.8e-41;

Matches 434; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

Qy 126 GGGGTTGAAGCGGCACATCAACAACAACCTGGGCAACCGCTTACGAGTTCCTGAGAC 185
Db 3 GGCTGAGAAAGCAGACGACGCGGCTGTGAAGATCGACACTTACGTGCTGGGGGAC 62

QY	186	CGTGGGCAAGGGGACCTTACGGGAAAGGTGAAGAGGACGAGAGAGCT---	CGGGGGGCT	24.2
Db	63	CTGGGCGTGGGACCTTGGGCAAGGAAAGTTGGAAACATCAATTGACAGGCCATTA	12.22	
QY	243	GGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAAATGAGCAGGATCTGTGCAT	30.20	
Db	123	ACTGGCAGTTAAGATCTTAAATATGACAGAAAGATTGCGAGTTAGATGTTGGAAAAAT	18.22	
QY	303	ACGAGAGGAGATTGAGATCATGTCTTCACTCAACACACCCCAATCATTTGCCATTCATGA	36.2	
Db	183	AAAAACAGAAATTCAAAATCTTAACTCTTCCGTCACTCCATATTAATCAACCTACCA	24.22	
QY	363	AGGTGTTGAAATAGCAGCAAGATTTGATTTGATCATGGAGTATGCCAGCCGAGCATCT	42.22	
Db	243	AGTGATCAGCACTCCAAACAGACTTTTATATGTAATGAAATATATGTCTGGAGGTGAATT	30.20	
QY	423	GTATGATTATCATCAGTAGAGCGGCAACGGCTGATGAGCGGAGCCGACGACTTCTTCG	48.2	
Db	303	GTTCACATCACTCTGTAAACAGCGGAGGGTTGAAGAGTGAAGCTGCGGGCTTTCCA	36.2	
QY	483	ACAGATCGTCTGCTGCCCTGSCACTACTGCCACAGAACGGAGTGTTCACCGAGATCTCAA	54.2	
Db	363	GCAGATTCTGTGCGCCGTGGACTACTGTCAACAGGACATATGTTGTTCACAGGACCTGAA	42.22	
QY	543	GGTGAAGAAACATCCTCTTATAGATGCCAATGGAACATCAAGATTGTGACTTTGGCCCTTC	60.2	
Db	423	GCACAGAAACGGTGTGTGCTGAGCGCCACAGATGAATGTCTAAATATGCTGACTTGGACTTC	48.2	
QY	603	CAACCTGTACCAACAAAGCAAGTTCCTCCAGACGTTCTGTGGAGACCCCTCTACGCGCTC	66.2	
Db	483	TAAATATGATGTCAAGATGTGTGAATTTTACGAACCTAGCTGTGGAATCGCCAAATTAATGACG	54.2	
QY	663	GCCTGAGATATGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGAACAAGTGTCTTGGG	72.2	
Db	543	ACCGAGAGTCAATCTCAGAAAGGCTGATGTAGCGGCTCTGAGATGTGAATCTGGAGCTGTGG	60.2	
QY	723	CGTTCTCTCTGTACATCTGTGTGATGGCACATATGCCCTTTGACGGGACGAGATCATTAAC	78.2	
Db	603	TGTATATCTGTATAGCCCTTCTCTGTGTGACACCCCTCCGGTTGAGAGATGAGACATGTCTAC	66.2	
QY	783	ACTGTGAAAGCAAAATCAGTAAACGGGGCTTAACGCTGAGCGGSCCAAGCGCTCCGATGCTG	84.2	
Db	663	GCTCTTTAAGAAAGATCCGAGGGGTGTGTTCTACATCCCGGAGATATCTCAACCGTTCTAT	72.2	
QY	843	TGGC---CTGATCCGGTGCCTTTAATGTGTGAACCCCAACCCGTCCGGGCCACTGTGAGGA	89.9	
Db	723	TGCCACTGTGTGATGCACATGTCTGACAGGTGACCCCTTGAAACGAGCACTATCAAAAGA	78.2	
QY	900	TGTAGCACTCATGTGTGGGTCAACTGGGGTTACCA 937		
Db	783	CATACGAGACATGAATGGTTTAAACAGATTTGGCCA 820		

Search completed: May 15, 2005, 22:58:52
Job time : 305.196 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 14, 2005, 22:57:45 ; Search time 985.386 Seconds
(without alignments)
11372.259 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 acgagagtcggtgagcttact.....gaatctgcacaaagctcagc 1893

Scoring table: IDENTITY NTC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	4 AAC90433	AAC90433 Murine Ly
2	1893	100.0	3073	13 ABD33083	ABD33083 Murine ca
3	1683.4	88.9	2929	6 AAD31710	Aad31710 Rat SNF1/
4	1360.2	71.9	1884	12 AD141462	Ad141462 Novel hum
5	1360.2	71.9	3353	6 ABK14000	Abk14000 cDNA enco
6	1360.2	71.9	3353	12 AD141460	Ad141460 Novel hum
7	1357.6	71.7	3360	6 AAD26459	Aad26459 Human kin
8	1357.6	71.7	3395	4 AAK94280	Aak94280 Human ful
9	1357	71.7	3395	12 ADJ30885	Adj30885 Full leng
10	1357	71.7	3404	13 ABD33085	Abd33085 Human can
11	1357	71.7	3463	12 ADJ96554	Adj96554 Human cal
12	1355.4	71.6	2043	6 ABZ11333	Abz11333 Human pol
13	1355.4	71.6	2501	12 ADM43851	Adm43851 Novel hum
14	1355.4	71.6	3443	5 ABX71420	Abx71420 Human cel
15	1355.4	71.6	3443	10 ADF76964	Adf76964 Novel hum
16	1355.4	71.6	3443	10 ADP81952	Adp81952 Leukemia
17	1355.4	70.7	3443	12 ADO20171	Ado20171 Human pro
18	1338.8	70.7	2291	4 AAF44659	Aaf44659 Novel pro
19	1338.8	70.7	2291	12 AD129357	Ad129357 Human MAR
20	1314.2	69.4	3200	4 AAF75338	Aaf75338 Human TGF

21	1230	65.0	1893	10 AB277163	Ab277163 Human pro
22	1059.4	56.0	37278	13 ABD33082	ABD33082 Murine ca
23	840.2	44.4	2616	6 ABQ72599	Abq72599 Human MDD
24	840.2	44.4	2619	6 ABQ72698	Abq72698 Human MDD
25	723	38.2	39699	13 ABD33084	ABD33084 Human can
26	506.4	26.8	2884	4 AAD14328	Aad14328 Human pro
27	506.4	26.8	6828	4 AAD14327	Aad14327 Human pro
28	506.4	26.8	6828	10 ADE38420	Ad38420 Human pro
29	506.4	26.8	6828	11 ADN95767	Adn95767 Human BRC
30	506.4	26.8	6828	12 ADJ74808	Adj74808 Marker ge
31	506.4	26.8	6828	12 ADL25352	Adl25352 Human ARK
32	506.4	26.8	6828	12 ADQ19733	Adq19733 Human sof
33	506.4	26.8	6828	12 ADP43253	Adp43253 Human pro
34	506.4	26.8	6828	12 ADT5680	Adt5680 Breast ca
35	506.4	26.8	6828	13 ADR25680	Adr25680 Human sof
36	356.8	18.8	3594	5 AAD03994	Aad03994 Human pro
37	348.8	18.4	587	4 AAK93286	Aak93286 Human CDN
38	348.8	18.4	587	4 AAK91887	Aak91887 Human CDN
39	348.8	18.4	587	12 ADL29723	Adl29723 5' end of
40	348.8	18.4	587	12 ADL28314	Adl28314 5' end of
41	296.8	15.7	1454	4 ABL21269	Ab121269 Drosophila
42	267	14.1	1723	4 AAH13802	Aah13802 Human CDN
43	236.6	12.5	1594	4 AAF44655	Aaf44655 Novel pro
44	236.6	12.5	1594	8 AAL60326	Aal60326 Mouse 207
45	236.6	12.5	1594	12 AD129353	Ad129353 Mouse MAR

ALIGNMENTS

RESULT 1	AAC90433	standard; cDNA; 2902 BP.
ID	AAC90433	
XX	AAC90433;	
XX	19-MAR-2001 (first entry)	
DE	Murine Lymph node Stromal cell kinase 1 coding sequence.	
XX		
KW	Murine; Lymph node Stromal cell kinase; M1SK-1; autoimmune disorder;	
KW	wound healing; periodontal disease; inflammatory disease; tumour;	
KW	infection; allergy; sg.	
XX		
OS	Mus musculus.	
XX		
PN	W0200073468-A1.	
XX		
PD	07-DEC-2000.	
XX		
PF	26-MAY-2000; 2000MO-US014696.	
XX		
PR	28-MAY-1999; 99US-0136781P.	
XX		
XX	(IMMUNEX CORP.	
PA	Bird TA, Varca GD, Martin U, Anderson DM;	
XX	WPI, 2001-061546/07.	
DR	P-PSDB; AAB50056.	
XX		
PT	Novel murine and human kinase nucleic acids useful for treating	
PT	inflammations, infections, tumors, allergies, autoimmune diseases, and	
PT	for stimulating or suppressing immune responses.	
XX		
PS	Claim 1; Page 86-87; 106pp; English.	
XX		
CC	The present sequence is the coding sequence for Murine Lymph node Stromal	
CC	cell kinase 1 (M1SK-1). The protein encoded by the present sequence is	
CC	useful for treating a variety of disorders listed in the disclosure of	
CC	the specification, including autoimmune disorders, allergic reactions,	
CC	myeloid or lymphoid cell deficiencies, wound healing and tissue repair	
CC	and replacement, burns, incisions and ulcers, periodontal disease,	

CC Inflammatory diseases, tumours and bacterial, viral or fungal infection
XX Sequence 2902 BP; 654 A; 794 C; 796 G; 658 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 1893; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGTGGTGGCTTACTTCCAGCGCCGAGCCAGGCTCCCTGGGCTTCCGCTGGCC 60
DB 123 ATGAGTGGTGGCTTACTTCCAGCGCCGAGCCAGGCTCCCTGGGCTTCCGCTGGCC 182
QY 61 TCGGAGCGCCGCGCGCTGGCGAGCGGCTCATCAAGTGGCTAAACCTTGATGAAG 120
DB 183 TCGGAGCGCCGCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAAACCTTGATGAAG 242
QY 121 AAGCAGCGGTGAAGCGGCACATCAACAAACAGACCTGGCGACCGCTACGATTCCTG 180
DB 243 AAGCAGCGGTGAAGCGGCACATCAACAAACAGACCTGGCGACCGCTACGATTCCTG 302
QY 181 GAGACCTGGGCAAGGCGACCTTACGGGAAGTGAAGAGCAGAGAGCTGGGGCGT 240
DB 303 GAGACCTGGGCAAGGCGACCTTACGGGAAGTGAAGAGCAGAGAGCTGGGGCGT 362
QY 241 CTGGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAGATGAGCAGATCTTGCGAC 300
DB 363 CTGGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAGATGAGCAGATCTTGCGAC 422
QY 301 ATACGAGGAGATGAGATCATGTCTTCACTCAACACCCGACATCATTTGCCATCAT 360
DB 423 ATACGAGGAGATGAGATCATGTCTTCACTCAACACCCGACATCATTTGCCATCAT 482
QY 361 GAAGTGTGAGAAATGACAGCAAGATTGTGATGTCAATGAGATGCGACCGAGCGAT 420
DB 483 GAAGTGTGAGAAATGACAGCAAGATTGTGATGTCAATGAGATGCGACCGAGCGAT 542
QY 421 CTGTATGATTAATCATGATGAGCGGCAAGCTGAGAGCGGAGACCGCAGGCAATTTCTC 480
DB 543 CTGTATGATTAATCATGATGAGCGGCAAGCTGAGAGCGGAGACCGCAGGCAATTTCTC 602
QY 481 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGAACGGATCGTTCAACGAGATCTC 540
DB 603 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGAACGGATCGTTCAACGAGATCTC 662
QY 541 AAGCTGAAAAACATCTCTTCTAGATGCCAATGAAACATCAAGATGCTGACTTTGGCCTC 600
DB 663 AAGCTGAAAAACATCTCTTCTAGATGCCAATGAAACATCAAGATGCTGACTTTGGCCTC 722
QY 601 TCCAACTGTACCAAGGCAAGTCTCTCAGACGCTCTGTGGGAGCCCTCTCTACGCGC 660
DB 723 TCCAACTGTACCAAGGCAAGTCTCTCAGACGCTCTGTGTGGAGCCCTCTCTACGCGC 782
QY 661 TCGCCTGAGATGATCAACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGTCTCTG 720
DB 783 TCGCCTGAGATGATCAACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGTCTCTG 842
QY 721 GGGCTTCTCTGTACATCTGTGTGATGCGACCATGCTCTTTGACGGGCGAGATCAATAA 780
DB 843 GGGCTTCTCTGTACATCTGTGTGATGCGACCATGCTCTTTGACGGGCGAGATCAATAA 902
QY 781 ACACTGTGAAGCAATCAAGTACGGGGCTTACCGTGAAGCGCGCCAAAGCGCTCCGATGCC 840
DB 903 ACACTGTGAAGCAATCAAGTACGGGGCTTACCGTGAAGCGCGCCAAAGCGCTCCGATGCC 962
QY 841 TGTGGCTGATCCGCTGGCTGTATATGTGAACCCCAACCGGTGGGGCAACATGAGAGAT 900
DB 963 TGTGGCTGATCCGCTGGCTGTATATGTGAACCCCAACCGGTGGGGCAACATGAGAGAT 1022
QY 901 GTAGCCAGTCATTTGTGGGTCAACTGGGGTTACACCAACCGGATCGGGGAAACAGAAAGCC 960
DB 1023 GTAGCCAGTCATTTGTGGGTCAACTGGGGTTACACCAACCGGATCGGGGAAACAGAAAGCC 1082
QY 961 CTGGCGTGAAGGTGGGCAACCTGATGTGACTTTGGCGGGCTTCATGGCGGACGTGGTTA 1020

DB 1083 CTGGCGTGAAGGTGGGCAACCTGATGTGACTTTGGCGGGCTTCATGGCGGACGTGGTTA 1142
QY 1021 CGTGGCTCTCGCGCCCTCTCTGAGAAATGAGCCAGAGTGTGACCTTCTTCAAGCAG 1080
DB 1143 CGTGGCTCTCGCGCCCTCTCTGAGAAATGAGCCAGAGTGTGACCTTCTTCAAGCAG 1202
QY 1081 CAGTGTCCGGAGGTGGAAGCACTGTACTTGGGCTGAGACCGGCAATCTCTTAAAGAG 1140
DB 1203 CAGTGTCCGGAGGTGGAAGCACTGTACTTGGGCTGAGACCGGCAATCTCTTAAAGAG 1262
QY 1141 TCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGAGACCCGCTAGAGTACTCT 1200
DB 1263 TCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGAGACCCGCTAGAGTACTCT 1322
QY 1201 TCTGGCTCTGGCAAGCAGCCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1260
DB 1323 TCTGGCTCTGGCAAGCAGCCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1382
QY 1261 ACTTGTCAAGGGAGGTACAGAGACCTTCAAGAACTCAAGACCGGTGCTGATCTCA 1320
DB 1383 ACTTGTCAAGGGAGGTACAGAGACCTTCAAGAACTCAAGACCGGTGCTGATCTCA 1442
QY 1321 GGGCAGCTGTCCCTGTGATCCCTGCTCCAAAGGAGGCAATCCTTAAAGTCTGCA 1380
DB 1443 GGGCAGCTGTCCCTGTGATCCCTGCTCCAAAGGAGGCAATCCTTAAAGTCTGCA 1502
QY 1381 CAGCGTGAATCTGTATTACTCTCTCCAGAGCCAGAGATCTTGGGAACTTTAGAC 1440
DB 1503 CAGCGTGAATCTGTATTACTCTCTCCAGAGCCAGAGATCTTGGGAACTTTAGAC 1562
QY 1441 GCCAGTATGTGTTGTGATGAGGAGCCCGGTGAGCAGAGTCTTCAACAGCTTCAAGG 1500
DB 1563 GCCAGTATGTGTTGTGATGAGGAGCCCGGTGAGCAGAGTCTTCAACAGCTTCAAGG 1622
QY 1501 CTCTCTCTCAACCGGCAAGGCAATCTCAAACTCAATGGCAATTCCTCCGCAAGCCTTA 1560
DB 1623 CTCTCTCTCAACCGGCAAGGCAATCTCAAACTCAATGGCAATTCCTCCGCAAGCCTTA 1682
QY 1561 GAAGCATTACCCCTAGCAGCTTTGGCTCTCCAGACCAACTGGGCTCTCCATCTCTGA 1620
DB 1683 GAAGCATTACCCCTAGCAGCTTTGGCTCTCCAGACCAACTGGGCTCTCCATCTCTGA 1742
QY 1621 GCCCGGCCAGCCGCCCTCAGGGGCTGTGATGAGACAGCATCTGTCTTCCAGATCC 1680
DB 1743 GCCCGGCCAGCCGCCCTCAGGGGCTGTGATGAGACAGCATCTGTCTTCCAGATCC 1802
QY 1681 TTTGACCAATTGGAATGTGCTGAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1740
DB 1803 TTTGACCAATTGGAATGTGCTGAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1862
QY 1741 GTGACAACTGAGGGGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAAGCATGTGGCAG 1800
DB 1863 GTGACAACTGAGGGGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAAGCATGTGGCAG 1922
QY 1801 GAATCTTGGGGAATGACTGCTTTTCTGACAGACTGCAAGAGGTGACTGACGCTTAC 1860
DB 1923 GAATCTTGGGGAATGACTGCTTTTCTGACAGACTGCAAGAGGTGACTGACGCTTAC 1982
QY 1861 AGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893
DB 1983 AGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 2015

RESULT 2
ABD33083
ID ABD33083 standard; cDNA; 3073 BP.
XX
AC ABD33083;
AC
XX
DT 18-NOV-2004 (first entry)
XX
DE Murine cancer-associated (CA) cDNA M07-004.

XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW ss; cancer; cystostatic.
 XX Mus musculus.
 OS WO2004058146-A2.
 PN 15-JUL-2004.
 PD 15-DEC-2003; 2003WO-US040081.
 PF 17-DEC-2002; 2002US-00322281.
 PR (SAGR-) SAGRES DISCOVERY INC.
 XX PI Morris DW, Malandro MS;
 XX WPI; 2004-499109/47.
 DR P-PSDB; AB084756.
 XX Novel human cancer associated protein encoded within open reading frame
 PT of cancer associated gene, useful as targets for diagnosing cancer.
 PS Disclosure; SEQ ID NO 22; 182zp; English.
 XX The invention relates to cancer-associated proteins (CAP) and the cancer-
 CC associated (CA) nucleic acids encoding them. The invention also relates
 CC to a method for treating cancers involving administering to a patient an
 CC inhibitor of CAP, and a method of screening for anticancer activity in a
 CC potential drug involving providing a cell that expresses a CA gene,
 CC contacting a tissue sample derived from a cancer cell with an anticancer
 CC drug candidate and monitoring the effect of the anticancer drug candidate
 CC on expression of the CA gene. The CAP proteins are useful for detecting
 CC cancer associated with expression of a CAP protein in a test cell sample
 CC and for screening for a bioactive agent capable of modulating the
 CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 CC cancer, involving determining the expression of a CA nucleic acid in a
 CC tissue. This sequence represents murine CA cDNA of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3073 BP; 690 A; 827 C; 844 G; 712 T; 0 U; 0 Other;

Query Match 100.0%; Score 1893; DB 13; Length 3073;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAATCGGTGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 60
 DB 92 ATGGAATCGGTGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 151
 QY 61 TCGGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCGCTAAACCTTGATGAAG 120
 DB 152 TCGGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCGCTAAACCTTGATGAAG 211
 QY 121 AAGCAGCGGTGAGAGCGGACCATCAAAACAACTGCGGACCGCTACGAGTTCCTG 180
 DB 212 AAGCAGCGGTGAGAGCGGACCATCAAAACAACTGCGGACCGCTACGAGTTCCTG 271
 QY 181 GAGACGCTGGGCAAGGCACTACGAGAGGTGAAGAGCAAGAGACTGGGGCGT 240
 DB 272 GAGACGCTGGGCAAGGCACTACGAGAGGTGAAGAGCAAGAGACTGGGGCGT 331
 QY 241 CTGGTGGCCATCAAGTCATCAGAAAGAAATCAAAATGATGAGAGTTCGTCGAC 300
 DB 332 CTGGTGGCCATCAAGTCATCAGAAAGAAATCAAAATGATGAGAGTTCGTCGAC 391
 QY 301 ATACGAGGAGATGATGATGATGCTTCACTCAACCAACCCCAATGATTCATCAT 360
 DB 392 ATACGAGGAGATGATGATGATGCTTCACTCAACCAACCCCAATGATTCATCAT 451

QY 361 GAAAGTTTGAGAAATGACGAAAGATTGATTTGTATGAGATATGCGACCGAGCGAT 420
 DB 452 GAAAGTTTGAGAAATGACGAAAGATTGATTTGTATGAGATATGCGACCGAGCGAT 511
 QY 421 CTGTATGATTAATCATCATGAGGCGGCAAGGCTGAGTGAAGCGGCAAGCAATTTCTTC 480
 DB 512 CTGTATGATTAATCATCATGAGGCGGCAAGGCTGAGTGAAGCGGCAAGCAATTTCTTC 571
 QY 481 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGAGCGGATTCGACCGAGATCTC 540
 DB 572 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGAGCGGATTCGACCGAGATCTC 631
 QY 541 AAGCTGAAAACATCTCTTCTAGATGCCAATGAAACATCAAGATTGCTGACTTGGCTTC 600
 DB 632 AAGCTGAAAACATCTCTTCTAGATGCCAATGAAACATCAAGATTGCTGACTTGGCTTC 691
 QY 601 TCCAACTGTATCAACAAAGGCAAGTTCCTCCAGACGTTCTGTGAGGACCTCTCTACGCC 660
 DB 692 TCCAACTGTATCAACAAAGGCAAGTTCCTCCAGACGTTCTGTGAGGACCTCTCTACGCC 751
 QY 661 TCGCCTGAGATAGTCAACGAGGAAAGCCCTATATGCGGCGCAGAGGTGACAGCTGCTCTG 720
 DB 752 TCGCCTGAGATAGTCAACGAGGAAAGCCCTATATGCGGCGCAGAGGTGACAGCTGCTCTG 811
 QY 721 GAGCTTCTCTGTATCAATCTGTATGATGCAATGCAATCCCTTTTGAAGGAGATCAATAA 780
 DB 812 GAGCTTCTCTGTATCAATCTGTATGATGCAATGCAATCCCTTTTGAAGGAGATCAATAA 871
 QY 781 ACACTGTGAGAAATCAATGATGACGAGGCTTACCGTGAAGCGGCCCAAGCTGCGATGCC 840
 DB 872 ACACTGTGAGAAATCAATGATGACGAGGCTTACCGTGAAGCGGCCCAAGCTGCGATGCC 931
 QY 841 TGTGGCTGATATCCGGGCTGTTAATGATGTAACCCCAACCGGTGGGCACTGAGAGAT 900
 DB 932 TGTGGCTGATATCCGGGCTGTTAATGATGTAACCCCAACCGGTGGGCACTGAGAGAT 991
 QY 901 GTAGCAGATCATTTGTGGGTCAACTGGGTTTACCAACCGAGTGGGGAACAGAAAGCC 960
 DB 992 GTAGCAGATCATTTGTGGGTCAACTGGGTTTACCAACCGAGTGGGGAACAGAAAGCC 1051
 QY 961 CTGCTGAGAGGTGGGCAACCTTATGATGTAATTTGGCGGAGCTTCATGCGGACTGTGTA 1020
 DB 1052 CTGCTGAGAGGTGGGCAACCTTATGATGTAATTTGGCGGAGCTTCATGCGGACTGTGTA 1111
 QY 1021 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGAGAGCTTCTTCAAGCAG 1080
 DB 1112 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGAGAGCTTCTTCAAGCAG 1171
 QY 1081 CAGTGCCTGGAGGTGAAAGCACTGTACTGAGCTGAGCGGCAACATTTCTTAAAGAG 1140
 DB 1172 CAGTGCCTGGAGGTGAAAGCACTGTACTGAGCTGAGCGGCAACATTTCTTAAAGAG 1231
 QY 1141 TCCGAAAGAGAAATGATGATGCTCAAAATCTGCAAGGTGACCCGCTGAGATACCTCT 1200
 DB 1232 TCCGAAAGAGAAATGATGATGCTCAAAATCTGCAAGGTGACCCGCTGAGATACCTCT 1291
 QY 1201 TCTGGCTCTGGCAAGGCAAGCTTTAAGCTTCGAAAGGCAATTTCTCAAGAAAAAGTCTCT 1260
 DB 1292 TCTGGCTCTGGCAAGGCAAGCTTTAAGCTTCGAAAGGCAATTTCTCAAGAAAAAGTCTCT 1351
 QY 1261 ACCTGCTCAGGGAGGATGACGAGAGACCTTCAAGAACTGACACCGGCTGATATCTCA 1320
 DB 1352 ACCTGCTCAGGGAGGATGACGAGAGACCTTCAAGAACTGACACCGGCTGATATCTCA 1411
 QY 1321 GGGCAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAAAGGCAATCTTAAAGATCTCGA 1380
 DB 1412 GGGCAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAAAGGCAATCTTAAAGATCTCGA 1471
 QY 1381 CAGCGTGAATCTGTGTTATCTCTCTCTCAAGACCCAGAGAGTCTGGGAACTTTAGAC 1440
 DB 1472 CAGCGTGAATCTGTGTTATCTCTCTCTCAAGACCCAGAGAGTCTGGGAACTTTAGAC 1531
 QY 1441 GCCAGTGAATGTGTTTGAAGTGGGAAACCGGTGAGAGAGAAATCTCCACAGGCTTCAGGG 1500

Db 1532 GCCAGGATGTTGTTGATGGGGACCCGCTGGAGAGAACTCTCCACAGGCTTCAGGG 1591
Qy 1501 CTCTCTCTCCACCGCAAGGCGATTTCTCAAACTCAATGGCAAGTTCTCCGCAAGCTTAA 1560
Db 1592 CTCTCTCTCCACCGCAAGGCGATTTCTCAAACTCAATGGCAAGTTCTCCGCAAGCTTAA 1651
Qy 1561 GAAGGACATACCCCTGACACTTTTGCTCCCTGGACCAACTGGCTCTCCCAATCCCTGCA 1620
Db 1652 GAAGGACATACCCCTGACACTTTTGCTCCCTGGACCAACTGGCTCTCCCAATCCCTGCA 1711
Qy 1621 GCCCGGCCCAAGCCGCTCAGGCGCTGTGAGTGAAGACAGACTCTGCTCCGAGTCC 1680
Db 1712 GCCCGGCCCAAGCCGCTCAGGCGCTGTGAGTGAAGACAGACTCTGCTCCGAGTCC 1771
Qy 1681 TTTGACCAATTGACATTTGCTGAAAGCTTTCCGAAAACCCCACTGAGGCGCTGTGCT 1740
Db 1772 TTTGACCAATTGACATTTGCTGAAAGCTTTCCGAAAACCCCACTGAGGCGCTGTGCT 1831
Qy 1741 GTGGAACAACCTGAGGGGGCTTGAAGAGCCCTCCCTGAGAAAGTCTGAAAGCGATGTGCGAG 1800
Db 1832 GTGGAACAACCTGAGGGGGCTTGAAGAGCCCTCCCTGAGAAAGTCTGAAAGCGATGTGCGAG 1891
Qy 1801 GAATCCTTGGGGGATGACTGCTTTCTCTGACAGACTGCGAAGAGTGACTGCAAGCTTAC 1860
Db 1892 GAATCCTTGGGGGATGACTGCTTTCTCTGACAGACTGCGAAGAGTGACTGCAAGCTTAC 1951
Qy 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893
Db 1952 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1984

RESULT 3

AAD31710
ID AAD31710 standard; cDNA; 2929 BP.AAD31710;
XX18-JUN-2002 (first entry)
XXRat SNF1/AMPK-Related Kinase (SNARK) cDNA.
XXRat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;
KM glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;
KM hyperglycaemic; drug screening; hypoglycaemia; ss.
XXRattus sp.
OSKey Location/Qualifiers
FH 83.1975
FT CDS /tag= a
PT /product= "Rat SNARK protein"
XXWO200212456-A2.
XX
PD 14-FEB-2002.
XXPF 02-AUG-2001; 2001WO-CA001109.
XX
PR 03-AUG-2000; 2000US-0222650P.
XX
PR 12-MAR-2001; 2001US-0274613P.
XX
PR 28-MAR-2001; 2001CA-02340783.
XX(ONBIO-) 1149336 ONTARIO INC.
XX
PA
PI Drucker DJ, Rosen CF, Lefebvre DL,
XX
DR MPI; 2002-241747/29.
XX
DR P-PSDB; AAE19885.
XXAMPK (AMP-activated protein kinase)-related kinase, designated SNARK
PT polypeptides and polynucleotides, useful for treating or preventing
PT diabetes, or other disorders of lipoprotein production leading toPT increased levels of cholesterol.
XXPS Example; Fig 2; 94pp; English.
XX

The invention relates to an AMPK (AMP-activated protein kinase)-related
CC kinase, designated SNARK polypeptides and polynucleotides. SNARK
CC (SNF1/AMP-activated protein kinase) is involved in stress response to
CC glucose deprivation. The polynucleotides are useful for expressing SNARK
CC protein in isolated form or as a protein conjugate. Activation of SNARK
CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
CC in other cell types such as heart and skeletal muscles, as well as
CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
CC to have insulin-like effects that would enhance the disposal of glucose
CC into muscle and reduce plasma glucose for the treatment of diabetes and
CC some type of disorders of lipoprotein production leading to increased
CC levels of cholesterol or triglycerides. SNARK or its variants may be
CC administered to a subject to treat or prevent a disease associated with
CC decreased expression of SNARK, such as diabetes. SNARK antibodies are
CC used to modulate SNARK activity either in vivo for therapeutic purposes,
CC or in vitro, for drug screening and related investigational purposes.
CC SNARK antagonists may be administered to increase fuel production,
CC decrease glucose uptake and increase levels of blood glucose in a patient
CC suffering from hypoglycaemia. The present sequence is rat SNARK cDNA
XX

SO Sequence 2929 BP; 650 A; 788 C; 791 G; 700 T; 0 U; 0 Other;

Query Match 88.9%; Score 1683.4; DB 6; Length 2929;
Best Local Similarity 93.4%; Pred. No. 0;

Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

Qy 1 ATGGAATCGGTGGCTTACTCCAGGCGCCGAGCGAGCTTCCTCGGCTCCGCGCTGGCC 60
Db 83 ATGGAATCGGTGGCTTACTCCAGGCGCCGAGCGAGCTTCCTCGGCTCCGCGCTGGCC 142
Qy 61 TCAGAGAGCGCCGCGCTGAGCGAGCGGCTCATCAAGTCGCTAACTCGATGAAG 120
Db 143 ACGGAGAGCGCCGCGCTGAGCGAGCGGCTCATCAAGTCGCTAACTCGATGAAG 202
Qy 121 AAGCAGGCGGTGAAGGCGGACCATCAAAACAACCTCGGCGAGCTTACGAGTTCCTG 180
Db 203 AAGCAGGCGGTGAAGGCGGACCATCAAAACAACCTCGGCGAGCTTACGAGTTCCTG 262
Qy 181 GAGACCTTGGGCAAGGCGCTTACGAGGAGGTGAAGAGCGACGAGAGCTTGGGCGCT 240
Db 263 GAGACCTTGGGCAAGGCGCTTACGAGGAGGTGAAGAGCGACGAGAGCTTGGGCGCC 322
Qy 241 CTGGTGGCCATCAAGTCCATCAAGGAAAGCAAAATCAAAAGATGAGAGATCGCTGCAC 300
Db 323 CTGGTGGCCATCAAGTCCATCAAGGAGGAAATCAAAAGATGAGAGATCGCTGCAC 382
Qy 301 ATACGAGGAGGATGAGATCAATGCTTCACTCAACCAACCCCATCATTTGGCATCAT 360
Db 383 ATACGAGGAGGATGAGATCAATGCTTCACTCAACCAACCCCATCATTTGGCATCAT 442
Qy 361 GAAGTCTTGAATGACGACAAATGTTGATGTCATGAGATGACAGCGAGCGAT 420
Db 443 GAAGTCTTGAATGACGACAAATGTTGATGTCATGAGATGACAGCGAGCGAT 502
Qy 421 CTGTATGATTATCATCAAGTGAAGGCGCAGCGCTGAGTGAAGCGGAGCGGCAATTTCTTC 480
Db 503 CTGTATGATTATCATCAAGTGAAGGCGCAGCGCTGAGTGAAGCGGAGCGGCAATTTCTTC 562
Qy 481 CGACAGATCGTGTGCGCTTGCATCACTGACCAAGAGCGGATGTTACCGAGATCTC 540
Db 563 CGACAGATCGTGTGCGCTTGCATCACTGACCAAGAGCGGATGTTACCGAGATCTC 622
Qy 541 AAGCTGGAAGAAATCTTTTATAGTCCAAATGGAACATCAATGATGCTGACTTTGGCTTC 600
Db 623 AAGCTGGAAGAAATCTTTTATAGTCCAAATGGAACATCAATGATGCTGACTTTGGCTTC 682
Qy 601 TCCAACTGTATCAAAAGGCAAGTTCCTCCAGAGATTTCTGTGGAGCCCTCTACAGCC 660
Db 683 TCCAACTGTATCAAAAGGCAAGTTCCTCCAGAGATTTCTGTGGAGCCCTCTATAGCC 742

QY 661 TCGCCTGAGTACTCAACGGGAAACCCCTATGTGGGCCAGAGGTGACAGCTGTCTCTG 720
DB 743 TACCTGAGATGCTCAACGGGAAACCCCTATGTGGGCCAGAGGTGACAGCTGTCTCTG 802
QY 721 GGGGTTCTCTGTACATCTGTGTGATGACCAATGCTCTTTGACGGGCGAGATCATAA 780
DB 803 GGGGTTCTCTGTACATCTGTGTGATGACCAATGCTCTTTGACGGGCGAGATCATAA 862
QY 781 ACACGTGTGAAGAATCAGTAACGGGCTTACCGTGTAGCGGCCCAAGCCGTCCGATGCC 840
DB 863 ACCGTGTGAAGAATCAGTAACGGGCTTACCGTGTAGCGGCCCAAGCCGTCTGATGCC 922
QY 841 TGTGGCTGATCCGCTGTGTGTAAATGTGAACCCCAACCCGTGGGCGACACTGAGAGAT 900
DB 923 TGTGGCTGATCCGCTGTGTGTAAATGTGAACCCCAACCCGTGGGCGACACTGAGAGAT 982
QY 901 GTAGCCAGTATGTTGTGTGTCAACTGTGGGTTACACACCGGAGTCCGGGAAACAGAAAGCC 960
DB 983 GTAGCCAGTATGTTGTGTGTCAACTGTGGGTTACACACCGGAGTCCGGGAAACAGAAAGCT 1042
QY 961 CTGCGTGAAGGGTGGGACCCCTAGTGTGACTTTGGCGGGGCTCCAGTGGGAGTGTGTA 1020
DB 1043 CTGCGTGAAGGGTGGGACCCCTAGTGTGACTTTGGCGGGGCTCCAGTGGGAGTGTGTA 1102
QY 1021 CGTGGCTCTCGGCGCCCTCTCTGAGATGAGCCAGGTTGTGACGCTTCTTCAAGCAG 1080
DB 1103 CGTGGCTCTCGGCGCCCTCTCTGAGATGAGCCAGGTTGTGAGCTTCTTCAAGCAG 1162
QY 1081 CAGGTCCGGGAGTGTGAGCACTGTACTCTGGCTGTAGGCGGCAACTTCTTTAGAGAG 1140
DB 1163 CAGGTCCGGGAGTGTGAGCACTGTACTCTGGCTGTAGGCGGCAACTTCTTTAGAGAG 1222
QY 1141 TCCCGAAAGAAATGATCATGTGCTCAAAATCTGCAAGGTGACCCGCTGTAGATCTCT 1200
DB 1223 TCCCGAAAGAAATGATCATGTGCTCAAAATCTGCAAGGTGACCCGCTGTAGATCTCT 1282
QY 1201 TCTGCGCTCTGAGAGAGCACTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGCTCTCT 1260
DB 1283 TCTGCGCTCTGAGAGAGCACTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGGCTCTCT 1342
QY 1261 ACCTGTCTGAGGAGGTGACAGAGAGCCCTCAGAGAACTCAGACCGGCTGTACTCCA 1320
DB 1343 CCTCATGCGGGAGGTGACAGAGAGCCCTCAGAGAACTCAGACCGGCTGTACTCCA 1402
QY 1321 GGGCAGCTCTCTCTGTGTATCCCTGTCTCCCAAGAAAGGCAATCTTAAAGATCTCCA 1380
DB 1403 GGGCAGCTCTCTCTGTGTATCCCTGTCTCCCAAGAAAGGCAATCTTAAAGATCTCCA 1462
QY 1381 CAGCGTGAATCTGTGTACTCTCTCTCCAGAGCCGAGGAGTCTGGGGAATCTTATGAC 1440
DB 1463 CAGCGTGAATCTGTGTACTCTCTCTCCAGAGCCGAGGAGTCTGGGGAATCTTATGAC 1522
QY 1441 GCCAGT 1500
DB 1523 GCGAGT 1582
QY 1501 CTCTCTCTCTGAGGAGGCAATCTCAAACTCAATGTGCAAGTTCTCCGCAAGCTTAA 1560
DB 1583 ---CGCTCTCATGCAAGGAGGCAATCTCAAACTCAATGTGCAAGTTCTCCGCAAGCTTAA 1639
QY 1561 GAAGGCACTACCCCTGAGCACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
DB 1640 GAAGGCACTACCCCTGAGCACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1699
QY 1621 GCCCGGCGCAGCGCCCTCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
DB 1700 GCCCGGCGCAGCGCTCTCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1759
QY 1681 TTTGACCAATTTGACTTGT 1740
DB 1760 TTTGACCAATTTGACTTGT 1819

QY 1741 GTGAGCACTGTGAGGGGCTTTGAGCAGCTCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG 1800
DB 1820 GTGAGCACTGTGAGGGGCTTTGAGCAGCTCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG 1879
QY 1801 GAATCTGTGGGAGATGCTGTCTTCTGACAGACTGCCAAGAGGTGACTGTGAGCCTTAC 1860
DB 1880 GAATCTGTGGGAGATGCTGTCTTCTGACAGACTGCCAAGAGGTGACTGTGAGCCTTAC 1939
QY 1861 AGACAAGCCCTAGGAATCTGCTCAAGCTGAGC 1893
DB 1940 AGACAAGCCCTAGGAATCTGCTGAGACTGAGC 1972

RESULT 4
ADL14162
ID ADL14162 standard; cDNA, 1884 BP.
XX AC ADL14162;
XX DT 17-JUN-2004 (first entry)
XX DE Novel human gene 3700 coding region.

KW cyostatic; cardiac; hypotensive; antianginal; osteopathic;
KW antiarthritic; antirheumatic; neuroprotective; antiinflammatory;
KW antipsoriasis; antiaesthetic; cardiovascular; vitruce; analgesic; CNS;
KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;
KW nephrotropic; antithyroid; dermatological; immunomodulator;
KW cell proliferation disorder; cell differentiation disorder;
KW brain disorder; platelet disorder; breast disorder; colon disorder;
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;
KW thyroid disorder; testes disorder; haematopoietic disorder;
KW pancreatic disorder; skeletal muscle disorder; skin disorder;
KW dermal disorder; bone metabolism disorder; immune disorder;
KW inflammatory disorder; cardiovascular disorder;
KW endometrial cell disorder; liver disorder; viral disease; pain disorder;
KW metabolic disorder; neurological disorder;
KW central nervous system disorder; erythroid disorder;
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
KW cell proliferation; cell differentiation; cell growth; cell division;
human; gene; ss.

XX OS Homo sapiens.
XX XX
XX PN US2004058355-A1.
XX XX
XX PD 25-MAR-2004.
XX PF 25-APR-2003; 2003US-00423543.
XX XX
XX PR 30-SEP-1998; 98US-00163821.
XX PR 27-JAN-1999; 99US-0117580P.
XX PR 25-MAR-1999; 99US-00276400.
XX PR 30-JUL-1999; 99US-00365162.
XX PR 09-SEP-1999; 99US-00392189.
XX PR 05-OCT-1999; 99US-00412210.
XX PR 23-NOV-1999; 99US-00448076.
XX PR 29-FEB-2000; 2000US-0186061P.
XX PR 28-APR-2000; 2000US-0200688P.
XX PR 19-MAY-2000; 2000US-0205447P.
XX PR 30-JUN-2000; 2000US-00608921.
XX PR 31-JUL-2000; 2000US-0221925P.
XX PR 25-SEP-2000; 2000US-0234922P.
XX PR 25-SEP-2000; 2000US-0235035P.
XX PR 08-NOV-2000; 2000US-0246669P.
XX PR 09-NOV-2000; 2000US-00711216.
XX PR 14-NOV-2000; 2000US-0248325P.
XX PR 15-NOV-2000; 2000US-0248893P.
XX PR 22-DEC-2000; 2000US-0257511P.
XX PR 05-JAN-2001; 2001US-0260166P.

28-FEB-2001; 2001US-00797039.
 PR 27-FEB-2001; 2001US-00845044.
 PR 20-UTL-2001; 2001US-00909743.
 PR 31-JUL-2001; 2001US-00920346.
 PR 13-AUG-2001; 2001US-00928531.
 PR 14-AUG-2001; 2001US-00929218.
 PR 15-AUG-2001; 2001US-0312539P.
 PR 25-SEP-2001; 2001US-00963159.
 PR 08-NOV-2001; 2001US-0008016.
 PR 13-NOV-2001; 2001US-00012055.
 PR 15-NOV-2001; 2001US-00003690.
 PR 30-JAN-2002; 2002US-00060763.
 PR 25-MAR-2002; 2002US-00105989.
 PR 12-APR-2002; 2002US-00121911.
 PR 12-APR-2002; 2002US-00217168.
 PR 22-OCT-2002; 2002US-00278036.
 PR 02-JAN-2003; 2003US-00336489.
 PR 03-JAN-2003; 2003US-00336153.
 XX
 (MILL-) MILLENNIUM PHARM INC.
 PI Kapeller-Liebermann R, Hunter J, Meyers RE, Rudolph-Owen LA;
 PI Curtis RAJ, Olandt PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;
 PI Silos-Santiago I, Bandaru R;
 XX
 DR MPI: 2004-268788/25.
 DR P-PSDB; ADL4161.
 XX
 PT New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,
 PT 26176, 26343, 56638, 18610, 33217, 21967, h1983, 38555 or 593
 PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,
 PT heart failure and angina.
 XX
 PS Claim 1; SEQ ID NO 45; 139pp; English.
 XX
 XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,
 CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
 CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising
 CC any one of 40 nucleotide sequences (I). The nucleic acid molecules and
 CC polypeptides are useful for diagnosing and treating a subject having a
 CC disorder, or a subject at risk of developing a disorder, which is
 CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
 CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,
 CC m1983, 38555 or 593 activity, such as cellular proliferative and/or
 CC differentiative disorders, brain disorders, platelet disorders, breast
 CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
 CC ovarian disorders, prostate disorders, cervical disorders, spleen
 CC disorders, thymus disorders, thyroid disorders, testes disorders,
 CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
 CC skin (dermal) disorders, disorders associated with bone metabolism,
 CC immune, e.g. inflammatory disorders, cardiovascular disorders,
 CC endothelial cell disorders, liver disorders, viral diseases, pain
 CC disorders, metabolic disorders, neurological or central nervous system
 CC disorders, erythroid disorders, blood vessel disorders or angiogenic
 CC disorders (all claimed), e.g. cancer, heart failure, hypertension, Crohn's
 CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
 CC disease, psoriasis, or asthma. The nucleic acid molecules and
 CC polypeptides are also useful as modulating agents in regulating a variety
 CC of cellular processes, e.g. cell proliferation, differentiation, growth and
 CC division. This sequence encodes a novel human protein of the invention.
 CC Note: The sequences given in the specification are also available in
 CC electronic format from
 CC ftp://seqdata.nebplc.gov/sequence.html?DocID=20040058355.
 CC
 XX
 XX Sequence 1884 BP; 403 A; 599 C; 548 G; 334 T; 0 U; 0 Other;
 SQ
 Query Match 71.9%; Score 1360.2; DB 12; Length 1884;
 Best Local Similarity 84.2%; Pred. No. 0;
 Matches 155; Conservative 0; Mismatches 273; Indels 21; Gaps 2;
 QY 47 CCTCCGCTGCTCGAGAGAGCGCGCGCTGAGGAGCGGCTCATCAATGCGCCCA 106
 DB 35 CCACTCCCTCGGCGGAGAGCTAGCCCGCGCTGAGGAGGAGGCTGATCAAGTCCGCCA 94

QY 107 AACCTGTATGAAGAGAGCGCGGTGAAGCGGACCATCAACAACCTTGGGACC 166
 DB 95 AGCCCTTAATGAAGAGAGCGGCGGTGAAGCGGACCAACAAGCAACCTTGGGACC 154
 QY 167 GCTAAGAGTTCCTGAGAGCGCGGTGAAGAGGAGGCTTAACGGAGAGTGAAGAGCGAG 226
 DB 155 GCTAAGAGTTCCTGAGAGCGCGGTGAAGAGGAGGCTTAACGGAGAGTGAAGAGCGAG 214
 QY 227 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 286
 DB 215 AGAGCTCGGGGCGCGGTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 274
 QY 287 AGGATCTGTGCAATATGAGGAGGAGGAGATGATCATGTCTTCAACCAACCCGAC 346
 DB 275 AAGATCTGATGCAATATGAGGAGGAGGAGATGATCATGTCTTCAACCAACCCGAC 334
 QY 347 TCATTGCCATCCATGAGAGTGTGAAGAAATGAGCAAGATTTGATTTGATGATGATG 406
 DB 335 TCATTGCCATCCATGAGAGTGTGAAGAAATGAGCAAGATTTGATTTGATGATGATG 394
 QY 407 CCAAGCGAGGCGATCTGTATGATTTATCATCATGAGCGGCGAGCTGATGAGCGGAGC 466
 DB 395 CCAAGCGGCGGCGATCTGTATGATTTATCATCATGAGCGGCGAGCTGATGAGCGGAGC 454
 QY 467 CCAAGCGATTTCTTCCGAGAGATGCTGTGCTGCTGCACTAGTCCACAGAGCGGATCG 526
 DB 455 CTAGGCAATTTCTTCCGAGAGATGCTGTGCTGCTGCACTAGTCCACAGAGCGGATCG 514
 QY 527 TTCACGGAATCTCAAGCTGAGAAATCTCTTATATGCAATGAGAACTTCAATG 586
 DB 515 TCACCGAATCTCAAGCTGAGAAATCTCTTATATGCAATGAGAACTTCAATG 574
 QY 587 CTGACTTTGAGCTCTCAACCTGTACCAAAAGGCAAGTTCCGAGCGTTGAGG 646
 DB 575 CTGACTTTGAGCTCTCAACCTGTACCAAAAGGCAAGTTCCGAGCGTTGAGG 634
 QY 647 GCGCTTCTAAGCGCTGCTGAGATAGTCAAGGAGCGGAGCTTATGAGGCGGAGAGTGG 706
 DB 635 GCGCTTCTAAGCGCTGCTGAGATAGTCAAGGAGCGGAGCTTATGAGGCGGAGAGTGG 694
 QY 707 AAGCTGTGCTCTGCGGCTTCTCTGTATCATCTGTGATGAGCAATGCTTGAAG 766
 DB 695 AAGCTGTGCTCTGCGGCTTCTCTGTATCATCTGTGATGAGCAATGCTTGAAG 754
 QY 767 GCGAGGATATATAACATCTGTGAAGCAATCATGTAACGAGGCTTACCGTGAAGCGGCA 826
 DB 755 GCGATGATATATAACATCTGTGAAGCAATCATGTAACGAGGCTTACCGTGAAGCGGCA 814
 QY 827 AGCGTCCGATGCTGTGAGCTGTATCCGCTGCTGTAATGTGAACCCACCGCTGAG 886
 DB 815 AACCTGTATGCTGTGAGCTGTATCCGCTGCTGTAATGTGAACCCACCGCTGAG 874
 QY 887 CCAACTGAGATGTAAGCAATCATGTTGGGTCAATCGGGGTTTACCAACCGAGTGG 946
 DB 875 CCAACTGAGATGTAAGCAATCATGTTGGGTCAATCGGGGTTTACCAACCGAGTGG 934
 QY 947 GGAACAGGAGCGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006
 DB 935 GGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 994
 QY 1007 TGCGGATGTTAGTCTGCTCTGCGCGCGGCTCTCTGAGAGATGAGGAGGAGGAGTGA 1066
 DB 995 TGCGGATGTTAGTCTGCGCGCGGCTCTCTGAGAGATGAGGAGGAGGAGGAGTGA 1054
 QY 1067 GCTTTCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1126
 DB 1055 GCTTTCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
 QY 1127 ATTCTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1186
 DB 1115 ATTCTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1174

Oy		1187	CTGAGGAATACCTCTTCTGTGGCCCTGGCAAGAGAGCCTTAAGTTTCGAAAGGCATTCTCA	1246
Db		1175	CTGATGACACTGCCATCGCCCTGGCAAGACCACTCAAGCTGCCAAGGGCATTTCTCA	1234
Oy		1247	AGAAAAGTCTCTAACCCTGTCAAGGGAGAGTACAAGAGAACCCTCAGAACTCAGACCGG	1306
Db		1235	AGAGAAAGTGTCAGACTCTGTCAAGAAAGGGGTATACAGAGSACCCTCCGAGCTCAGCCAA	1294
Oy		1307	TGCTGTATACTCAGAGGAGACCTGTGCTCTGTATTCCTTGTCTCCAAAGAAAGGCATCC	1366
Db		1295	TCCCTGCACAGCCCAAGGGCAGGGCTGTCCC-----CCCTGTCTCCCAAGAGGGCATTC	1349
Oy		1367	TTAAGAACTTCGACAGGGTGATATGTGTTACTACTCTCTCCAGAGGCCGCGAGCTCG	1426
Db		1346	TCAGAAACCCCCGACAGCGCAGAGTGTGCTACTCTCTCTCCGAGCCCAAGTGAATTCG	1405
Oy		1427	GGAAGACTTTAGACCGCACAGTATGTGTTTGTCAGTGGGAGCCCGCTGAGAGCAAGTCTC	1486
Db		1406	GGAGACTTTTGGACCCAGGCCACAGTGTTTGTGATGGGGGATCCAGAGAGCAAGACCTC	1465
Oy		1487	CACAGGCTTCAAGGGCTCTCTCTCCACCGCAAGGGCATTCTCAAATCAATGSCAAATTCT	1546
Db		1466	CGAAGCTTTCAGGGCTCTCTCTCCATCGCAAGGCATCCTCAAACTCAATGGCAAGTTCT	1525
Oy		1547	CCCGACAGCCTTTAAGAGCACTAACCTTAGACACTTTGGCTCTCTCTGAGACCAATTGGCT	1606
Db		1526	CCCAAGACGCTTTAGAGAGCTCGCGGCCCCACCACTTCGGCTCTCTGTGATGAATCGGCC	1585
Oy		1607	CCTCCCATCTCGACGCGCGGCCGACGCGCCCTCAGGGGCTGTGAGTGAAGACAGCATCC	1666
Db		1586	CACCTCGCCCCCTGTGCCCCGGGCGACGCCAACCCTCAGGGGCTGTGAGCGAGACAGCATCC	1645
Oy		1667	TGTCCTCCGAGTCTTTGACCAATTGGACTTGCTGAAAGTCTTCCGAAAACCCCACTGA	1726
Db		1646	TGTCTCTGAGTCTTTGACCAAGCTTGACCTTGAAAGGCTGTGAGCGAGACAGCATCC	1705
Oy		1727	GGGGCTGTGTCTGTGTGACCACTCAGAGGGGCTTGAGCAGCTCTCTCTCAAG-----	1788
Db		1706	GGGGCTGTGTGTCTGTGTGACCAACTCAGGGGCTTGAGAGAGCCCCCTCAGAGGGCCCTG	1765
Oy		1781	-----GTCTGAAGCGATGTGTGGCAGGAATCCTTGGGGGATATGATGCTTTCTCTGACAG	1834
Db		1766	GAACTGCTCTGAGGCGCTGGCGGCGAGATCCTTTGGGGGACAGCTGCTTTTCCTTGACAG	1825
Oy		1835	ACTGCCAAGAGGTGACTCAGGCTTACAGACCAAGCCCTTAGAATCTGCTCAAAAGCTCAGC	1893
Db		1826	ACTGCCAAGAGGTGACAGGACCTTACGACAGGCACTAAGGGTCTGCTCAAAAGCTCACC	1884
<hr/>				
RESULT 5				
ABK14000 standard; cDNA, 3353 BP.				
ID	ABK14000			
XX	ABK14000;			
AC				
XX	02-JUL-2002 (first entry)			
DT				
XX				
DE	cDNA encoding human protein kinase 3700.			
XX				
KW	Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;			
KW	cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;			
KW	tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;			
KW	cell proliferation disorder; cell differentiation disorder; carcinoma;			
KW	haematopoietic neoplastic disorder; metastatic disorder; leukaemia;			
KX	cytostatic; antineuroleptic; gene; ss.			
OS	Homo sapiens.			
XX				
PH	Key			
FT	Location/Qualifiers			
CD	157..2043			
FT	/tag= a			
FT	/product= "Protein kinase 3700"			
FT	/note= "The coding region not including the terminator			

XX		(codon) is specifically claimed in claims 7 and 28"
PT		
XX	WO200224921-A2.	
EN		
XX	28-MAR-2002.	
PD		
XX		
PF	25-SEP-2001; 2001WO-US030115.	
XX		
PR	25-SEP-2000; 2000US-0234922P.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
P1	Curtis RAJ, Galvin KM;	
DR	WPI; 2002-352007/38.	
DR	P-PsDB; AAU79652.	
PT	Use of modulators of activity of 3700 protein for making medicament for	
PT	e.g., modulating protein phosphorylation or cell signalling, or for	
PT	treating or preventing cellular proliferative and/or differentiative	
PT	disorders.	
PS		
PS	Claim 28; Fig 1; 115bp; English.	
XX		
CC	The present invention relates to the isolation of a novel human protein	
CC	kinase designated 3700, and the polynucleotide sequence encoding it. The	
CC	invention also describes the use of a modulator of the activity of	
CC	protein kinase (PK) 3700 for making a medicament or pharmaceutical	
CC	composition for modulating the ability of a cell to phosphorylate an	
CC	amino acid residue of a substrate protein. Modulators of protein kinase	
CC	3700 activity are useful for modulating protein phosphorylation, cell	
CC	signalling, tumorigenesis, mitogenesis, transcription of a gene,	
CC	angiogenesis, tissue repair, tissue regeneration, establishment or	
CC	progression of atherosclerosis, and signalling across the blood-brain	
CC	barrier. The polynucleotide and polypeptide molecules for protein kinase	
CC	3700 may be used as diagnostic targets and therapeutic agents for	
CC	prognosticating, diagnosing, preventing, inhibiting, alleviating, or	
CC	curing PK-related disorders and cellular proliferative and/or	
CC	differentiative disorders (e.g. haematopoietic neoplastic disorders,	
CC	carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700	
CC	polynucleotide sequence can be used to express protein kinase 3700, to	
CC	detect a genetic alteration in a 3700 gene, in chromosome mapping, for	
CC	tissue typing, in forensic biology, and as surrogate markers. The present	
CC	sequence encodes human protein kinase 3700	
CC		
SQ	Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;	
XX		
Query Match	71.9%; Score 1360.2; DB 6; Length 3353;	
Best Local Similarity	84.2%; Pred. No. 0;	
Matches 1565; Conservative	0; Mismatches 273; Indels 21; Gaps 2	
Gy	47 CCTCGGCGCCTTGAGGACGGCCGCGTGGCGGAGCGGGCTCATCAATGCCTTA 106	
Dbb	191 CCATCTCCCTCGGCGGAGACTAGCCGCCGTGGCGGAAGGCTGATCAATGGCCA 250	
Gy	107 AACCTTGATGAAGAAGCAGCGCGTGAAGCGGACCATCACAAACAACCTGCGGACC 166	
Dbb	251 AGCCGCTAATGAAGAAGCAGCGGCGGAGAACCAACAAGCAACCTGCGGACCC 310	
Gy	167 GCTACGAATTCTTGAAGCGCTGGGCAAGGCACTTAAGGGAAGTGAAGAAGGCAACG 226	
Dbb	311 GCTACGAATTCTTGAAGCGCTGGGCAAGGCACTTAAGGGAAGTGAAGAAGGCGGG 370	
Gy	227 AGAGTCGGGGCGTGTGGTGGCCATCAATGCATCAGGAAAGAAAATCAAAGATGAGC 286	
Dbb	371 AGAGTCGGGGCGGCTGTGGTGGCCATCAATGCATCAGGAAAGCAAAATCAAAGATGAGC 430	
Gy	287 AGAGTCTGCTCAATACGAGAGGAGATTAAGATCATGTCTTCACTCAACCAACCCCACA 346	
Dbb	431 AAGATCTGATCAATACGAGAGGAGATTAAGATCATGTCTCACTCAACCAACCCCTACA 490	
Gy	347 TCATTGCCATCATCAAGTTTGAAGATTAGCAGCAAGATTGTATTTGTCATGAGATATG 406	

Db 491 TCATTGCCATCATGAAGTGTGAGAAAGCAGCAAGATCGTATGCTCATGAGATATG 550
QY 407 CCAGCCGAGGCGATCTGTATGATTAATCACTGAGCGCGCAAGGCTGATGAGCGGAGC 466
Db 551 CCAGCCGAGGCGGACCTTATGATCACTGAGCGCGCAGCAGCTCATGAGGCGGAG 610
QY 467 CCAGGACATTTCTTCCGACAGATGATGCTGCGCCGCACTACTGCAACAGGAGATCG 526
Db 611 CTAGGACATTTCTTCCGACAGATGATGCTGCGCCGCACTACTGCAACAGGAGATCG 670
QY 527 TTCAACGAGATCTCAAGCTGAAAACATCTCTTAGATGCAATGAAAATCAAGATTTG 586
Db 671 TCACCGAGATCTCAAGCTGAAAACATCTCTTAGATGCAATGAAAATCAAGATTTG 730
QY 587 CTGACTTTGGCTCTTCCAACTGTATCCAAAGGCAAGTTCTTCAGACGTTCTGTGGA 646
Db 731 CTGACTTTGGCTCTTCCAACTGTATCCAAAGGCAAGTTCTTCAGACGTTCTGTGGA 790
QY 647 GCCCTCTATAGGCTGCGCTGAGATGATCAACGAGGAGCCCTATATGAGGCGCAGAGTGG 706
Db 791 GCCCTCTATAGGCTGCGCAGAGATGATCAAGGAGCCCTATCAAGGCGCAGAGTGG 850
QY 707 ACAGCTGATCTCTGAGGCGTCTCTGTATCACTCTGTGATGAGCACTAGCCCTTTGAGC 766
Db 851 ACAGCTGATCTCTGAGGCGTCTCTGTATCACTCTGTGATGAGCACTAGCCCTTTGAGT 910
QY 767 GGGAGATCATTAACACTGTGTAAAGAAATCAAGTAAAGGCGGCTTACCGTGAAGCGGCA 826
Db 911 GGGATACCAATTAAGATCTGTAGTAAAGATCAAGTAAAGGCGGCTTACCGGAGGCACTTA 970
QY 827 AGCGCTCCATGCTGCTGAGCTGATCCGAGGCTGTTAATGTTGAACCCCAACCCCTGGG 886
Db 971 AACCTCTGATGCTGCTGAGCTGATCCGAGGCTGTTAATGTTGAACCCCAACCCCTGGG 1030
QY 887 CCACACTGAGAGATGATAGCAGTCAATGTGTGATCACTGAGGTTAACAACCGAGATGG 946
Db 1031 CCACCGTGAAGATGATGAGCAGTCACTGTGTGATCACTGAGGTTAACAACCGAGATGG 1090
QY 947 GGGAAACAGAAAGCCCTGCTGAGAGGTGGGCAACCTTGTGTGATCTTGGCCGAGCTTCA 1006
Db 1091 GAGAGAGAGAGGCTCCGCAATGAGGTGGGCAACCTTGTGTGATCTTGGCCGAGCTTCA 1150
QY 1007 TGCGGAGTGGTAAAGCTGCTGCTGCGGCGCCCTCCCTGAGATATGAGCCCAAGGTGCA 1066
Db 1151 TGCGTACGTGGCTCGGCGCTTCTCCGCGCCCTCTCTGAGAAATGGGCGCAAGGTGCA 1210
QY 1067 GCTTCTTCAAGCAGCATGTGCGCGGAGGTGAAAGCATGTAACCTGGGCTGAGCGGCAAC 1126
Db 1211 GCTTCTTCAAGCAGCATGTGCGCGGAGGTGAAAGCATGTAACCTGGGCTGAGCGGCAAC 1270
QY 1127 ATTCTTTAAGAAATCCCGAAAGAGATGATGCTCAAAATCTTGCAAGGTGACCCGG 1186
Db 1271 ATTCTTTAAGAAATCCCGCAAGAGATGATGCTCAAAATCTTGCAAGGTGACCCGG 1330
QY 1187 CTGAGGATGCTCTTCTGCGCCCTGCAAGGAGCCTTAAAGCTTCCGAAAGGCAATTTCA 1246
Db 1331 CTGATTAACCTGCTCAATGCGCTTGGCAAGGCAACCTCAAGCTCCAAAGGGCAATTTCA 1390
QY 1247 AGAAAAAGTCTTAACTCTGCTGAGGAGGTGAGAGAGAGCCTCAGAGAACTCAGACCGG 1306
Db 1391 AGAAGAGAGGTGAGGCTCTGCAAGAGGAGGTACAGAGAGCCTTCGAGACTCAGGCCAA 1450
QY 1307 TGCCTGATATCTCAAGGAGAGCCTGTCTGCTGTATCTCTGCTCCAAAGGAAAGGATCC 1366
Db 1451 TGCCTGAGAGCCAGGAGGAGGCTGCGC-----CCCTGCTCCCAAGAGGAGATTC 1501
QY 1367 TTAAGAGGCTGAGCAGCGGTGAATCTGTATTAATCTCTTCCAGAGCCCAAGGATCTG 1426
Db 1502 TCAGAGAGCCCGCAGAGCGGAGATCTGCTATCACTCTCTTCCAGGCCCAAGGAAATCTG 1561
QY 1427 GGGAACTTTAAGACGCAAGTATGTGTTGTGAGTGGGAGCCCGGTGGAGCAGAAATCTC 1486
Db 1562 GGGAGCTCTTGAACGAGCGAGCAGTGTGTTGTGAGTGGGAAATCCCAAGAGCAGAAAGCTTC 1621

QY 1487 CACAGGCTTCAAGGCTTCTCTCTCCACCGAAGGCAATTTCAAACTCAATGGCAAGTTCT 1546
Db 1622 CGCAAGCTTCAAGGCTTCTCTCTCCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1681
QY 1547 CCCGACAGGCTTAAAGAGCACTACCCCTAGACACTTTGGCTCTCCAGAACATGCGCT 1606
Db 1682 CCCAGAGGCTTAAAGAGCTGCGGCGCCCAACCACTTGGCTCTCCAGAACATGCGCT 1741
QY 1607 CCTCCATCTGAGAGCGGCGCCAGCGCCCTCAGGAGGCTGAGAGAGAGCATCC 1666
Db 1742 CACTGCGCCCTGCGCCCGGCGCAGCGACCTTCAAGGAGCTGAGAGAGAGAGCATCC 1801
QY 1667 TGTCTTCCAGATCTTTGACCAATTTGACCTTTGCTGAAACGCTTTCCGAAACCCCACTGA 1726
Db 1802 TGTCTTCCAGATCTTTGACCAATTTGACCTTTGCTGAAACGCTTTCCGAAACCCCACTGC 1861
QY 1727 GGGGCTGTGTGTGTGTGAGACAACTGAGAGGAGGCTTGAAGAGCTCTCTCAAG----- 1780
Db 1862 GGGGCTGTGTGTGTGTGTGAGACAACTGAGAGGAGGCTTGAAGAGGCTCTCTCAAGAGGCTCTG 1821
QY 1781 -----GTCTGAAGCGATGAGTGGCAGAGAAATCTTGGGAGATGAGCTGCTTTCTGTGAG 1834
Db 1922 GAAGCTGCTGAGAGGCTGCGCGCAGAGATCTTTGGGAGACAGCTGCTTTTCCCTGAGAG 1881
QY 1835 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAAGCAAGCCCTAAGAAATCTGCTCAAGCTGAGC 1893
Db 1982 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAAGCAAGCTTGAAGAGGCTGCTCAAGCTGAGC 2040

RESULT 6
ADL14160
ID ADL14160 standard; cDNA; 3353 BP.
XX
AC ADL14160;
XX
DT 17-JUN-2004 (first entry)
XX
DE Novel human gene 3700 cDNA.
XX
KW cytosolic; cardiac; hypotensive; antianginal; osteoporotic;
XX antiarthritis; antirheumatic; neuroprotective; antiinflammatory;
XX antidiabetic; antidiabetic; cardiovascular; virucide; analgesic;
XX angiotensin inhibitor; angiotensin stimulator; cerebroprotective;
XX nephrotropic; antihypertensive; dermatological; immunomodulator;
XX cell proliferation disorder; cell differentiation disorder;
XX brain disorder; platelet disorder; breast disorder; colon disorder;
XX kidney disorder; renal disorder; lung disorder; ovarian disorder;
XX prostate disorder; cervical disorder; spleen disorder; thymus disorder;
XX thyroid disorder; testes disorder; haematopoietic disorder;
XX pancreatic disorder; skeletal muscle disorder; skin disorder;
XX dermal disorder; bone metabolism disorder; immune disorder;
XX inflammatory disorder; cardiovascular disorder;
XX endometrial cell disorder; liver disorder; viral disease; pain disorder;
KW metabolic disorder; neurological disorder;
KW central nervous system disorder; erythroid disorder;
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
KW cell proliferation; cell differentiation; cell growth; cell division;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN US2004058355-A1.
XX
PD 25-MAR-2004.
XX
PF 25-APR-2003; 2003US-00423543.
XX
PR 30-SEP-1998; 98US-00163821.
PR 27-JUN-1999; 99US-01175800.
PR 25-MAR-1999; 99US-00276400.

XX Human full-length cDNA, SEQ ID NO: 2918.
DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX (HELI-) HELIX RES INST.
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI, 2001-524255/58.
XX P-PSDB; AAM93360.
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX Claim 8; SEQ ID NO 2918; 1380bp + Sequence listing; English.
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
XX Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
SQ
Query Match 71.7%; Score 1357; DB 4; Length 3395;
Best Local Similarity 84.1%; Pired. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCGCGCGCTGGCGTGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAATGCCCTTA 106
DB 152 CCACTCCCTCGCGCGAGAGCTAGCCCGCGCTGGCGGAGGGCTGATCAATCCGCCCA 211
QY 107 AACCTGTGTAAGAAAGACAGCGCGTGAAGCGGACCATCAAAACAACCTGCGGACCC 166
DB 212 ACCCCCTAATGAAGAGAGCGCGGTGAAGCGGACCAACAAGACCAACCTCGGACCC 271
QY 167 GCTAGAGGTTCTCTGGAGAGCGTGGGCAAGGCGCTTAAGGGAAGTGAAGAGGCAAG 226
DB 272 GCTACGAGTTCCTGGAGAGCGCTGGGCAAGGCGCTTAAGGGAAGTGAAGAGGCGCG 331
QY 227 AGAGGCTCGGGGGGTCTGGTGGCCATCAATGCTCATGAGAAAGCAAAATCAAAAGTAA 286
DB 332 AGAGGCTCGGGGGGTCTGGTGGCCATCAATGCTCATGAGAAAGCAAAATCAAAAGTAA 391
QY 287 AGAGTCTGTGACATACGAGAGGAGATTGAGATCAATGCTCATCAACCAACCCGACAC 346
DB 392 AGAGTCTGTGACATACGAGAGGAGATTGAGATCAATGCTCATCAACCAACCCGACAC 451
QY 347 TCATTCGATCCATGAAGTGTGTTGAGAAATGACAGCAAGATGTGATGTTCATGAGTATG 406
DB 452 TCATTCGATCCATGAAGTGTGTTGAGAAATGACAGCAAGATGTGATGTTCATGAGTATG 511

QY 407 CCAGCCGAGCGATCTGTATGATTACATGATGAGCGGCCACCGGCTGATGAGCGGACG 466
DB 512 CCAGCGGGGGCGACTTTATGATCAATCAAGCGGAGCGAGCTCAAGAGCGCGAAG 571
QY 467 CCAGGATTTCTTCCGACAGATGCTGTGCTGCGCTGACCTACTGCTCCACAGAGGGATG 526
DB 572 CTAGGATTTCTTCCGAGAGATGCTCTGCGCTGACCTATGCTCATCAGAACAGATTG 631
QY 527 TTCACGAGATCTCAAGCTGGAAGAACATCTTGAATGCGAATGGAACATCAAGATTG 586
DB 632 TCACCGAGATCTCAAGCTGGAAGAACATCTTGAATGCGAATGGAACATCAAGATTG 691
QY 587 CTGACTTTGGCCCTTCCAACTGTACCAAAAGGCAAGTTCTCCAGAGCTTCTGTGGGA 646
DB 692 CTGACTTTGGCCCTTCCAACTGTACCAAAAGGCAAGTTCTCCAGAGCAATTTGTGGGA 751
QY 647 GCCCTCTTACGCGCTGCGCTGAGATGTCAGACGGGAAGCCCTATGTTGGGCGCAGAGTGG 706
DB 752 GCCCCTCTATGCTGCTGCGCAGAGATTGTCAATGGAAAGCCCTACAGAGCCGAGAGTGG 811
QY 707 ACAGCTGTCTCTGGGCGTTTCTCTGTACATCTGTGTCATGAGCACATGCCCTTTGACG 766
DB 812 ACAGCTGTCTCTGGGCGTTTCTCTGTACATCTGTGTCATGAGCACATGCCCTTTGATG 871
QY 767 GCGAGGATCAATAAAGACTGTGAAGCAAAATCAATCAAGGCGCTTACCGTGAACCGCCCA 826
DB 872 GGCATACCAATAAAGACTGTGAAGCAAAATCAATCAAGGCGCTTACCGGAGCCACCTTA 931
QY 827 AGCCGTCCGATGCGCTGTGCGCTGATCCGGTGGCTGTAAATGGAAGCCCAACCGTCCGG 886
DB 932 AACCTCTGATGCTGTGCGCTGATCCGGTGGCTGTAAATGGAAGCCCAACCGTCCGG 991
QY 887 CCACTGTGAAGATGTAGCAAGTCAATGTGTGCTCAATGAGGTTTACCAACCGGAGTGC 946
DB 992 CCACTGTGAAGATGTAGCAAGTCAATGTGTGCTCAATGAGGTTTACCAACCGGAGTGC 1051
QY 947 GGAAGAAGAAAGCCCTGAGAGGTTGGGAGCACTTGTGTGATCTTTGGCGCGGCTTCCA 1006
DB 1052 GAAAGAGGAAGGCTTCGATGAGGTTGGGAGCACTTGTGTGATCTTTGGCGCGGCTTCCA 1111
QY 1007 TGGCGGACTGTGTTACGTGCTCTCGCGGCGCCCTCTCTGGAAGATGAGGCCAAGGTGCA 1066
DB 1112 TGGCTGATCTGCTCGCGGCTTCTCTCGCGGCGCCCTCTCTGGAAGATGAGGCCAAGGTGCA 1171
QY 1067 GCTTTCTTAAAGACAGCAAGTCCGAGAGTGAAGCACTGTACTGTGAGCTGAGCGGCAAC 1126
DB 1172 GCTTTCTTAAAGACAGCAAGTCCGAGAGTGAAGCACTGTACTGTGAGCTGAGCGGCAAC 1231
QY 1127 ATTCTCTTAAAGATCCGAAAGGAAATGACATGTGCTCAAAATCTGCAAGTGAACCGG 1186
DB 1232 ATTCTCTTAAAGATCCGAAAGGAAATGACATGTGCTCAAAATCTGCAAGTGAACCGG 1291
QY 1187 CTGAGGATACCTTCTTCTGCGCGTGAAGAGCGCTTAAAGCTTCCGAAGGATCTCA 1246
DB 1292 CTGATATACCTGCTCATCTGCTTGGGCAAGGCACTTCAAGCTGCGAAGGAGATCTCA 1351
QY 1247 AGAAAAAGTCTCTTACTTCTGTGAGGAGGATCAAGAGGACCTTCAAGAACTCAAGACCG 1306
DB 1352 AGAAGAAAGGTGTCAGCTCTGCAAGAAAGGGTACAGAGGACCTTCCGAGCTCAGCCCA 1411
QY 1307 TGCCTGATATCTCAAGGAGAGCTGTCTCTGTGTATCTCTGCTCCCAAGAAAGGATCC 1366
DB 1412 TCCCTGCGAGCCCAAGGAGGCGCTGCC-----CCCTGTCTCCCAAGAAAGGATTC 1462
QY 1367 TTAAGAAAGTCTGAGACAGCGTGAATCTGCTTCACTCCCTCTCAAGGCCCAAGAGTCTG 1426
DB 1463 TCAAGAAAGCCCGAAGAGCGAGTCTGCTACTACTCTCTCCGAGCCCAAGTAATCTG 1522
QY 1427 GGAAGCTTTTGAAGCGCAGTGTATGTTTGTGAGTGGGAGCCCGTGAAGCAAGTCTC 1486
DB 1523 GGAAGCTTTTGAAGCGCAGTGTATGTTTGTGAGTGGGATCCCAAGAGAGCAAGAGCTTC 1582
QY 1487 CAAGGCTTCAGGGCTCTCTCTTCAACCGCAAGGCAATCTCAAACTCAATGCAAGTTCT 1546

DB 1583 CGAAGCTTCAGGAGCTGCTCTCCATGCAAAAGCATCTCAAACTCAATGAGCAAGTTCT 1642
QY 1547 CCCGACAGGCTTGAAGAGCACTAACCCCTGACACTTTGGCTCCCTGACCAACTGAGCT 1606
DB 1643 CCCAGACAGCTTGAAGCTGCGGCGCCCACTGCTTGGCTCCCTGAGAACTGCGCC 1702
QY 1607 CCTCCATCTGACAGCCGCGCCAGCCCTCAGAGGCTGTGATGAGAGACAGATCC 1666
DB 1703 CACTGCGCCCTGCGCCCGGCGCAGCCGCACTCAGAGGCTGTGAGAGAGAGATCC 1762
QY 1667 TGTCTCCGAGCTCTTGAACCAATGGAATTGGAATTGCTGTAAGCTCTTCCGAAACCTCACTGA 1726
DB 1763 TGTCTCTGAGCTCTTGAACCAAGCTGGAATCTGCTGCAAGCGCTCCAGAGCGCCCACTGCG 1822
QY 1727 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCCCTCCCTCAAGAG----- 1780
DB 1823 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCGCCCTCCCTCAAGAGGGCCCTG 1882
QY 1781 -----GTCTGAAGCGATGTGTGAGAGAAATCTTGGGGATAGCTGCTTTCTGTGACAG 1834
DB 1883 GAAGCTGCTGAGGCGCTGCGCGAGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1942
QY 1835 ACTGCCAAGAGTGACTGACGCTTACAGACAAAGCCTTAGAATCTGCTCAAGTCAAGC 1893
DB 1943 ACTGCCAAGAGTGACAGGACCTTACCGACAGGCACTGAGGGTCTGCTCAAGTCAAGC 2001

RESULT 9

ADL30885
ID ADL30885 standard; cDNA; 3395 BP.

XX AC ADL30885;

XX DT 20-MAY-2004 (first entry)

XX DE Full length human cDNA clone Segid 2918.

XX KW human; medicine; signal transduction; glycoprotein; transcription;

XX KM oligo-capping method; ss; gene.

XX OS Homo sapiens.

XX PN BP1396543-A2.

XX PD 10-MAR-2004.

XX PF 07-JUL-2000; 2003BP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000UP-00118774.

XX PR 02-MAY-2000; 2000UP-00183865.

XX PR 07-JUL-2000; 2000BP-00114089.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR MPI; 2004-204755/20.

XX DR P-PSDB; ADL30886.

XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX PS length human cDNAs.

XX PS Example 1; SEQ ID NO 2918; 1340BP; English.

XX CC This invention relates to a novel primers useful for synthesizing full

XX CC length cDNA molecules that encode human proteins. Specifically, it refers

XX CC to secretory or membrane proteins that are potential therapeutic agents/

XX CC target molecules in the field of medicine, and in particular genes

XX CC encoding proteins that are associated with signal transduction,

XX CC glycoproteins and transcription. The present invention describes a method

CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX

SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;

Query Match 71.7%; Score 1357; DB 12; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

QY 47 CTTCCGCGCTGCGCTGAGAGAGCGCCGCGCTGCGGAGCGGCTCATGAAGCGCTCA 106
DB 152 CCACTCCCTCGCGCGAGACTAGCCCGCGCTGCGGAGGCTGATCAAGTGGCCA 211
QY 107 AACCTGATGAAG 166
DB 212 AGCCCTTATGAG 271
QY 167 GCTACGAGTTCTGAG 226
DB 272 GCTACGAGTTCTGAG 331
QY 227 AGAGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286
DB 332 AGAGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 287 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 346
DB 392 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
QY 347 TCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406
DB 452 TCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
QY 407 CCAAGCGAG 466
DB 512 CCAAGCGAG 571
QY 467 CCAAGCGAG 526
DB 572 CCAAGCGAG 631
QY 527 TTGACGAG 586
DB 632 TTGACGAG 691
QY 587 CTGACTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
DB 692 CTGACTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751
QY 647 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
DB 752 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
QY 707 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 766
DB 812 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871
QY 767 GGCAGATCATTAAG 826
DB 872 GGCAGATCATTAAG 931
QY 827 AGCCCTTATGAG 886
DB 932 AGCCCTTATGAG 991
QY 887 CCAAGCTGAG 946
DB 992 CCAAGCTGAG 1051
QY 947 GGCAG 1006
DB 1052 GGCAG 1111

PS Example 1; SEQ ID NO 11; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytoskeletal, neuroprotective, immunomodulator and anti-inflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
SQ Sequence 3463 BP; 732 A; 1057 C; 919 G; 755 T; 0 U; 0 Other;
Query Match 71.7%; Score 1357; DB 12; Length 3463;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCTCCGCGCTGGCTCGAGAGCGCCCGCGCTGGCGGACGGGCTCATAGTCGCTTA 106
Db CCACTCCCTCGCGCGAGAGCTAGCCGCGCTGGGGAAGGCTGATCAAGTCGCCCA 282
QY 107 AACCTCTGATGAAGACGACGCGGTGAAGCGGCATCATCAACACACACCTCGCGCAC 166
Db 283 AGCCCCCTAATGAAGACGACGCGGTGAAGCGGCACACCAACAGCAACCTCGCGCAC 342
QY 167 GCTACGAGTCTCGAGAGCGCTGGGCAAGGGCACCTACGGGAAGTGAAGAAGCGACGAG 226
Db 343 GCTACGAGTCTCGAGAGCGCTGGGCAAGGGCACCTACGGGAAGTGAAGAAGCGCGGG 402
QY 227 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 286
Db 403 AGAGCTCGGGCGCGCTGGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 462
QY 287 AGAGTCTGCTGCAATACGAGAGGAGATGAGATCATGTCTTCACTCAACCCGCCACA 346
Db 463 AATATCTGATGCAATACGAGAGGAGATGAGATCATGTCTTCACTCAACCCGCCACA 522
QY 347 TCATTCGATCCATGAAAGTGTGAGAAATGACGCAAGATTTGATGATGATGATGATG 406
Db 523 TCATTCGATCCATGAAAGTGTGAGAAATGACGCAAGATTTGATGATGATGATGATG 582
QY 407 CCAGCGAGGCGATCTGTATGATTACATCAGTAGCGGCGACGGCTGAGTGAAGCGGAGC 466
Db 583 CCAGCGGCGGCGACTTTATGACTACATCAGCGAGCGGCGACGCTGATGAGCGGAG 642
QY 467 CCAAGCAATTTCTTCCGACAGATGATGTCTGCGCTGCACTA CTGCGCACAGAGCGGATCG 526
Db 643 CTAGGATTTCTTCCGCGAATGATGTCTGCGCTGCACTA CTGCGCACAGAGGTTG 702
QY 527 TTCACCGAATCTCAGAGTGGAAAAATCTCTTCTAGTGGCAATGAAAAATCAAGATTG 586
Db 703 TTCACCGAATCTCAGAGTGGAAAAATCTCTTCTAGTGGCAATGAAAAATCAAGATTG 762
QY 587 CTGACTTTGGCTCTCTCAACCTGTACCAAGGCAAGTTCTCTCAGAGCTTCTGAGGA 646
Db 763 CTGACTTTGGCTCTCTCAACCTGTACCAAGGCAAGTTCTCTGAGCAACTCTCTGAGGA 822
QY 647 GCCCTTCTTACGCGCTGCGCTGAGATAGTCAACGGGAAGCCCTATGTTGGGCCAGAGGTG 706
Db 823 GCCCTTCTTACGCGCTGCGCGCAGAGATTTGTCATGAGGAAGCCCTACAGGCGCGAAGGTG 882
QY 707 AAGAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGATGAGCAATGCTCTTTGACG 766
Db 883 AAGAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGATGAGCAATGCTCTTTGATG 942
QY 767 GGCAGATCAATAAAGCACTGGTGAAGCAATCAGTAAACGGGGCTTACCGTGAACCGGCCA 826
Db 943 GGCATACATTAAGATCTTAGTAAACAGATCAGCAACGGGGCTTACCGGGAACCGCACTTA 1002

QY 827 AGCCGTCCGATGCGCTGTGCTGATTCGGTGTGCTGTATATGTGAACCCGACCGGTGGG 886
Db 1003 AACCTCTGATGCTGTGGCTGCTATCCGGGTGCTGTGATGTGAACCCGACCGCGGG 1062
QY 887 CCACACTGAGAGATGTAAACCAAGCATTTGGTGGTCAACTGGGGTTTCAACACCGGACTCG 946
Db 1063 CCACCTGAGAGATGTGGCCATGTACTGTGGGTCAACTGGGGTTTCAACCGCAACCGAGTGG 1122
QY 947 GGGAAACGAAAGCCCTGCGTGAAGGGTGGGACCTTGAATGTGACTTTGCGGGGCTTCA 1006
Db 1123 GAGAGCAGAGAGCTCCGCAATGAGAGTGGGACCTTGGCAATGACTCTGCGCGGCTTCA 1182
QY 1007 TGGCGAATGTATTACGTGTCTCTGCGCCCTCTCTGAGAAATGAGACCAAGTGTGCA 1066
Db 1183 TGGCTGACTGGCTCCGGCGTTTCTCTCCGCGCTCTCTGAGAAATGAGGCGCAAGTGTGCA 1242
QY 1067 GCTTCTTCAAGCAGCATGTCGCCGGGAGTGGAGCATCTGACTTGGGCTGAGACCGGAC 1126
Db 1243 GCTTCTTCAAGCAGCATGTCCTGTGGGGAGACCAACCTTGGCTGAGACCGGACG 1302
QY 1127 ATTCTCTTAAAGATCCCGAAAGAGATGACATGTGCTCAAAATCTGCAAGTGAACCGG 1186
Db 1303 ATTCTCTTAAAGATCCCGAAAGAGATGACATGTGCTCAAAATCTGCAAGTGAACCGG 1362
QY 1187 CTGAGATACCTTCTTCTGCGCTTGGCAAGACGCTTAAAGCTTCCGAAAGGCAATTTCA 1246
Db 1363 CTGATGACACTGTCCCATTCGCTTGGCAAGACGCAACTCAAGCTGCAAGAGGCAATTTCA 1422
QY 1247 AGAAAGATCTTACTCTGCTGACGGGAGGTACAGAGAGACCTTCAAGATCTGACACCGG 1306
Db 1423 AGAAAGATGTGATAGCTTCTGCAAGAGGAGGTACAGAGAGACCTTCCGAGCTTCAAGCCAA 1482
QY 1307 TGCCTATCTCCAGGGCAGCGCTGTCCCTGTGTATCCCTGTGCTCCCAAGGAAGGCAATC 1366
Db 1483 TCCCTGAGACCCAGGGCAGCGCTGTCCCTGTGTATCCCTGTGCTCCCAAGGAAGGCAATC 1533
QY 1367 TTAAGAGTCTGACAGCGGTGAATCTGTGTTACTCTCTCCAGAGCCGACGAGTCTG 1426
Db 1534 TCAAGAGCGCCGACAGCGGAGTCTGTGTTACTCTCTCCAGAGCCGACGAGTCTG 1593
QY 1427 GGAATCTTTAGACCGCAGTGAATGTGTTGTGATGAGGAGACCCGCTGAGCAGAACTTC 1486
Db 1594 GGAATCTTTAGACCGCAGTGAATGTGTTGTGATGAGGAGATCCCAAGGACCAAACTTC 1653
QY 1487 CACAGGCTTCAAGGGCTCTCTCTCCACGCAAGGGCAATTCCTCAAACTCAATGGCAAGTTC 1546
Db 1654 CCGAGCTTCAAGGGCTCTCTCTCCACGCAAGGGCAATTCCTCAAACTCAATGGCAAGTTC 1713
QY 1547 CCGGCAAGCTTTAGAGGCACTACCCCTAGCACTTTGGCTCTCTGAGCACTGAGCTT 1606
Db 1714 CCGGCAAGCTTTAGAGGCACTACCCCTAGCACTTTGGCTCTCTGAGCACTGAGCTT 1773
QY 1607 CTCTCCATCTGACAGCGCGCGCCAGCGCGCTCAGGGGCTGTGATGAGCAGCATTC 1666
Db 1774 CACCTTGGCGCTGTGGCGCGCGCGCGAGCGGACCTCAGGGGCTGTGATGAGCAGCATTC 1833
QY 1667 TGTCTCCGAGTCTTTTGA CCAATTGGAATGTGCTGAAGTCTTCCGAAACCCCACTGA 1726
Db 1834 TGTCTCTGAGTCTTTTGA CAGCTGAACTTGTGCTGAAGGCTTCCAGAGCCCCCACTG 1893
QY 1727 GGGGCTGTGTCTGTGACA CACTGAGGGGGCTTGAAGAGCTCTCTCAGAG----- 1780
Db 1894 GGGGCTGTGTGTGTGACA CAACTCAAGGGGCTTGAAGAGCCCTCAGAGGGGCGCTG 1953
QY 1781 -----GTCTGAAGCATGTGTGAGAAATCTTGGGGATGAGTCTTTTCTTGAAG 1834
Db 1954 GAACTGTCTGAGGCGCTGCGCGCAGAGATCTTTGGGGGACAGCTCTTTTCCCTGAAG 2013
QY 1835 ACTGCCAAGAGGATGAGCATGAGCTACAGCAAGCCCTAAGAAATCTGCTCAAGCTCAGC 1893
Db 2014 ACTGCCAAGAGGATGAGCATGAGCTACAGCAAGCCCTAAGAGGCTGCTCAAGCTCAGC 2072

RESULT 12
AB211333
ID AB211333 standard; cDNA; 2043 BP.
XX
AC AB211333;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 215.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW hemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritis; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002MO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RM, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drennec RT;
XX
DR WPI; 2002-759812/82.
XX
DR P-PSDB; ABP69116.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 215; 1012bp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB211119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP63843) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2043 BP; 418 A; 659 C; 592 G; 374 T; 0 U; 0 Other;

Query Match 71.6%; Score 1355.4; DB 6; Length 2043;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CTTCCGCGCTGCTGCGAAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCGGCTTA 106
DB 140 CCACTCCCTGCGCGAGAGCTAGCCCGCGCTGCGGAGAGGCGTATCAAGTCGCCA 199

QY 107 AACCTGATGAAGAGACAGCGCGTGAGCGGACCATCAACAACACAACTTGGGAC 166
DB 200 AGCCCTTAATGAAGAGACAGCGCGTGAGGAGGACCAACAAGCAAACTTGGGAC 259
QY 167 GCTACGAGTTCGAGAGAGCTGGGCAAGGGCACTTAAGGAAAGTGAAGAGGACGAG 226
DB 260 GCTACGAGTTCGAGAGAGCTGGGCAAGGGCACTTAAGGAAAGTGAAGAGGCGG 319
QY 227 AGAGCTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286
DB 320 AGAGCTCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 287 AGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 346
DB 380 AAGATCGATGACATACGAGAGAGTGAAGATGATGATGATGATGATGATGATGATG 439
QY 347 TCATTCGATCCATGAGAGTGTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
DB 440 TCATTCGATCCATGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
QY 407 CCAAGCGAGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 466
DB 500 CCAAGCGGCGGCGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 559
QY 467 CCAAGCATTTCTTCCGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526
DB 560 CTAGGATTTCTTCCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
QY 527 TTGACCGAGATCTCAAGCTGAGAAACATCTTCTGATGCGATGAGAAATCAAGATTG 586
DB 620 TCACCGAGATCTCAAGCTGAGAAACATCTTCTGATGCGATGAGAAATCAAGATTG 679
QY 587 CTGACCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
DB 680 CTGACCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739
QY 647 GCCCTCTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
DB 740 GCCCTCTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
QY 707 ACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
DB 800 ACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
QY 767 GGCAGATCATAAACCTGCTGAGAGCAATCAAGTAAAGGAGCTTACCTGAGCGGCCA 826
DB 860 GGCATGACATTAAGATCTGATGAAACAGATCAAGCAAGGAGCTTACCTGAGCGGCCA 919
QY 827 AGCGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 886
DB 920 AACCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
QY 887 CCAACATGAGAGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 980 CCAACCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1039
QY 947 GGAACAGAGAGAGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
DB 1040 GGAACAGAGAGAGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1099
QY 1007 TGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
DB 1100 TGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
QY 1067 GCTTCTTCAAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
DB 1160 GCTTCTTCAAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
QY 1127 ATTCTCTTAAGAGTCCCGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1186
DB 1220 ATTCTCTTAAGAGTCCCGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1279
QY 1187 CTGAGGATACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246

DB 1280 CTATGACACTGCCCATGCGCTGGCAAGCAACCTCAAGCTCCAAAGGCGCTTCTCA 1339
QY 1247 AGAAAAATGCTCTACTCTGTCAAGGGAGGTACAGAGAGCCCTCAGAACTCAGACGG 1306
DB 1340 AGAAGAAAGGTGACGCTCTGACAGAAAGGGGTACAGAGAGCCCTCGAGGTGAGCCCA 1399
QY 1307 TGGCTGATCTCCAGGGGACGCTGTCCCTGTGTATCCCTGTCTCCAAAGAAAGGCAATCC 1366
DB 1400 TCCCTGTGAGCCCAAGGCGAGCTGCC-----CCCTGTCTCCCAAGAGGCAATTC 1450
QY 1367 TTAAGAAGTCTGACAGCGTGAATCTGTACTACTCTCTCCAGAGCCCAAGGAGCTG 1426
DB 1451 TCAAGAGCCCCCAAGCGAGGTCTGGCTACTACTCTCTCCAGAGCCCAAGGAGCTG 1510
QY 1427 GGGAACTCTTGAAGCGCAGTATGTCTTGTGATGGGGAGCCCGTGAAGAGAGCTC 1486
DB 1511 GGGAGGCTTGGAGCGAGGCGAGCGTGTGTGTGATGGGGATCCCAAGAGAGAGAGCTC 1570
QY 1487 CAGAAGCTTCAAGGCTCTCTCTTCCACCGCAAGGAGATTTCAAACTCAATGGCAAGTTCT 1546
DB 1571 CGAAGCTTCAAGGCTGCTCTCTCATGCAAAAGGCAATCTCAAACTCAATGGCAAGTTCT 1630
QY 1547 CCGGACAGGCTTGAAGAGCACTAACCCCTAGACCTTTGGCTCCCTGAGCACTAGGCT 1606
DB 1631 CCGAGACAGCTTGAAGCTGCGGCGCCCAACACTTGGCTCCCTGATGAATGAGCTGCCC 1690
QY 1607 CTTCCCATCTTGAAGCGCGGCGCCAGCCCTCAGAGGGCTGTGATGAGAGACAGATCC 1666
DB 1691 CACTGCGCCCGCGGCGCGGCGAGCGAGCTCTCAGAGGGCTGTGAGAGAGAGAGATCC 1750
QY 1667 TGTCTCTCCAGTCTCTTGAACCAATTGGAATTTGCTGAACTCTTCCGAAACCCCACTGA 1726
DB 1751 TGTCTCTGATGCTCTTGAACAGCTGGAATCTGCTGAACCGCTCCAGAGCCCACTGCG 1810
QY 1727 GGGGCTGTGTCTGTGGAACAACCTGAGGGGCTTGAAGCGCTCCCTCAGAAAG----- 1780
DB 1811 GGGGCTGTGTGTCTGTGGAACAACCTGAGGGGCTTGAAGCGCGCCCTCAGAGGGCGCTG 1870
QY 1781 -----GTCTGAAGCGATGTGTGAGAGATCTTGTGGGAGTAGCTCTTTCTGTGACAG 1834
DB 1871 GAGCTGCTGAGGCGCTGCGGAGAGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1930
QY 1835 ACTGCCAAGAGGTGACTGACGCTTACAGACAGACGCTTGAAGATCTGCTCAAGCTCAGC 1893
DB 1931 ACTGCCAAGAGGTGACGACGACCTTACGAGACGACGACGAGGCTGCTCAAGCTCAGC 1989

RESULT 13
ADM43851
ID ADM43851 standard; cDNA; 2501 BP.
AC ADM43851;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #215.
XX
KM ss; gene; human; arginine-rich protein; cancer; inflammation;
XX genetic disorder.
XX Homo sapiens.
OS
PN US2004053250-A1.
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
XX 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095-
PR 20-AUG-2002; 2002US-00252551.
XX

PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
PI Tang YT, Xue A, Drmanac RT;
XX WPI; 2004-238579/22.
XX
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
XX Disclosure; SEQ ID NO 215; 51pp; English.
PS
XX
XX The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
SQ Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;
XX
XX
XX Query Match 71.6%; Score 1355.4; DB 12; Length 2501;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;
QY 47 CTTCCGCCCTGCGCTGAGAGCGCCGCGCTGCGGAGCGGCTCATCAATGCGCTTA 106
DB 140 CCACTCCCTCGCGCGCAGAGCTAGCCGCGCTGCGGAGGCGTGAATCAATGCGCCCA 199
QY 107 AACCTGATGAAGAGAGAGAGCGCGGTGAAGCGGACCATCAACAACAACCTGCGGACCC 166
DB 200 AGCCCTTAATGAAGAGAGAGCGCGGTGAAGCGGACCAACAAGCAACCTGCGGACCC 259
QY 167 GCTACAGATTTCTGAGAGCGCTGCGGAGAGGCACTACGAGGAGGTGAAGAGGCGAG 226
DB 260 GCTACAGATTTCTGAGAGCGCTGCGGAGAGGCACTACGAGGAGGTGAAGAGGCGG 319
QY 227 AGAGCTCGGGGCGTGTGTGCGCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 286
DB 320 AGAGCTCGGGGCGCTGTGTGCGCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 379
QY 287 AGAGTGTGTGCGCATCAAGGAGAGATGAATCATGTCTTCACTCAACCAACCCCAACA 346
DB 380 AAGATGTGAATGACATACGAGGAGAGATGAATCATGTCTTCACTCAACCAACCCCAACA 439
QY 347 TCATTGCCATCCATGAAGTGTGAGATAGCAGCAAGATTGATGATTCATGAGATATG 406
DB 440 TCATTGCCATCCATGAAGTGTGAGATAGCAGCAAGATTGATGATTCATGAGATATG 499
QY 407 CCAGCCGAGCGCATCTGTATGATTACATGAGGCGGCGCAGGCTGAGTGAAGCGGAGCG 466
DB 500 CCAGCCGAGCGCATCTTATGACTACATCAGGAGCGGCGCAGGCTGAGTGAAGCGGAGCG 559
QY 467 CCAGGATTTCTTCCAGAGATGTGTGTGCTGCGCTGCACTACTGCCCAAGAGGAGATG 526
DB 560 CTAGGCAATTTCTTCCGAGATGTGTGTGCTGCGCTGCACTACTGCCCAAGAGGAGATG 619
QY 527 TTCACCGAGATCTCAAGCTGGAAGAAATCATCTTCTAGATGCAATGGAATCATCAAGATTG 586
DB 620 TCACCGAGATCTCAAGCTGGAAGAAATCATCTTCTAGATGCAATGGAATCATCAAGATTG 679
QY 587 CTGACTTTGCGCTCTTCCAACTGTACACAAAGGCAAGTTCTTCCAGAGCTTGTGGGA 646
DB 680 CTGACTTTGCGCTCTTCCAACTGTACACAAAGGCAAGTTCTTCCAGAGCTTGTGGGA 729
QY 647 GCGCTTCTTACGCTGCTGAGATGTCAAGGAGAGCCCTTATGTGGGCGCGAGAGCTG 706
DB 740 GCGCTTCTTACGCTGCTGAGATGTCAAGGAGAGCCCTTATGTGGGCGCGAGAGCTG 799

QY 707 ACAGCTGCTCTGCGGCTTCTCTGTAACATCTGTGATGCGACCATGCGCTTTGACG 766
DB 800 ACAGCTGCTCTGCGGCTTCTCTGTAACATCTGTGATGCGACCATGCGCTTTGATG 859
QY 767 GGCAGATCATTAACACTGTGTAGAGCAATCAGTAACCGGGCTTACCTGAGCGGCCA 826
DB 860 GGCATGACCAATAGATCTTGTGTAAACAGATCAGCAACCGGGCTTACCGGAGCACTTA 919
QY 827 AGCGCTGATGCTGTGCGCTGATTCGGGTGCTGTAAATGATGTAACCCACCGGTGCGG 886
DB 920 AACCTTGTATGCTGTGCGCTGATTCGGGTGCTGTGTGTGTGTAAACCCACCGCGG 979
QY 887 CCAACTGAGAGATGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 980 CCAACCTGAGAGATGTGCGCAATGATGATGATGATGATGATGATGATGATGATGATG 1039
QY 947 GGGAAACAGAAACCTGCTGTGAGGGGTGCGACCTGATGATGATGATGATGATGATGATG 1006
DB 1040 GAGAGCAGAGAGCTCCGATGAGGGGTGCGACCTGATGATGATGATGATGATGATGATGATG 1099
QY 1007 TGGCGATGCTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
DB 1100 TGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
QY 1067 GCTTCTTCAAGCAGCAGTGTGCGCGGAGGTGGAACATGTAACCTGAGCTGAGCGGCAAC 1126
DB 1160 GCTTCTTCAAGCAGCAGTGTGCGCGGAGGTGGAACATGTAACCTGAGCTGAGCGGCAAC 1219
QY 1127 ATTCTCTTAAAGATCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1186
DB 1220 ATTCTCTTAAAGATCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1279
QY 1187 CTGAGATATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1246
DB 1280 CTGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1339
QY 1247 AGAAAGATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306
DB 1340 AGAAAGATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1399
QY 1307 TGCCTGATCTTCCAGGCGAGCTGTCTCTGCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366
DB 1400 TCCCTGAGCGCCAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
QY 1367 TTAAGAGTCTTCCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1426
DB 1451 TCAAGAGCGCCCGCAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1510
QY 1427 GGGAACTTAAAGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
DB 1511 GGGAGCTTGTGAGCGAGCGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1570
QY 1487 CACAGGCTTGAAGGCTCTCTCTCCAGCGCAAGGCAATCTCAAACTGATGCAAGTCTC 1546
DB 1571 CGCAGGCTTGAAGGCTCTCTCTCCAGCGCAAGGCAATCTCAAACTGATGCAAGTCTC 1630
QY 1547 CCGCAGAGCTTGAAGGCACTACCCCTGACACTTTGGCTCTCTGAGCAACTGAGCTC 1606
DB 1631 CCGCAGAGCTTGAAGGCACTACCCCTGACACTTTGGCTCTCTGAGCAACTGAGCTC 1690
QY 1607 CCTCCAGTCTGAGCG 1666
DB 1691 CACTTCCCGCTGAGCG 1750
QY 1667 TGTCTCTGAGTCTTGTGACCAATGTGACTTGTGCAACGTCTTCCGAAACCCCACTGA 1726
DB 1751 TGTCTCTGAGTCTTGTGACCAAGCTGTGACTTGTGCAACGTCTTCCGAAACCCCACTGA 1810
QY 1727 GGGGCTGTGTCTGTGTGACCACTGAGGGGGCTTGTGAGCACTCTCTCAGAAAG----- 1780
DB 1811 GGGGCTGTGTGTGTGTGACCACTGAGGGGGCTTGTGAGCACTCTCTCAGAAAGGGCGCTG 1870

QY 1781 -----GTCTGAAGCATGTGTGAGAGATCTTTGGGGGATAGTCTTTCTGTACAG 1834
DB 1871 GAAGCTGCTTGAAGCGCTGTGCGGAGATCTTTGGGGGACAGCTGTCTTCCCTGACAG 1930
QY 1835 ACTGCCAAGAGGTGAGCTGCGACCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
DB 1931 ACTGCCAAGAGGTGAGCTGCGACCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1989

RESULT 14
ABX71420
ID ABX71420 standard; cDNA; 3443 BP.
XX
XX AC ABX71420;
XX
XX DT 14-Apr-2003 (first entry)
XX
XX DE Human cell cycle-associated cDNA from clone DKFZpht3e3_7j3.
XX
XX KM Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200112659-A2.
XX
XX PD 22-FEB-2001.
XX
XX PE 18-AUG-2000; 2000MO-IB001496.
XX
XX PR 18-AUG-1999; 99US-014949P.
XX
XX PR 28-SEP-1999; 99US-0156503P.
XX
XX PA (GENU-) GERMAN HUMAN GENOME PROJECT.
XX
XX PI Wiemann S;
XX
XX DR WPI: 2001-327840/34.
XX
XX DR P-PSDB; ABUS3319.
XX
XX PT Nucleic acids having the sequences of clones isolated from libraries of
XX
XX PT different human tissues, useful in recombinant DNA methodologies.
XX
XX PS Claim 1; Page 942-943; 1095bp; English.
XX
XX CC This invention describes novel polynucleotides and polypeptides isolated
XX
XX CC from human cDNA libraries which can be used for gene therapy or in
XX
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by
XX
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX
XX CC associated with inappropriate polypeptide expression. The products of the
XX
XX CC invention may also be used to identify modulators of expression and
XX
XX CC activity and to down regulate expression and activity. The antibodies of
XX
XX CC the invention may also be used as diagnostic agents for detecting the
XX
XX CC presence of polypeptides in samples. This sequence encodes a polypeptide
XX
XX CC described in the disclosure of the invention

Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
SQ

Query Match 71.6%; Score 1355.4; DB 5; Length 3443;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;
QY 47 CTTCCGCGCTGGGCTGGAAGGCGCGCGCGCTGCGGAGCGGCTCATCAAGTCGCTTA 106
DB 162 CACTTCCCTGAGCGCGCGAGACTACCTCCGCGCTGCGGAGAGGCTATCAAGTCGCTTA 221
QY 107 AACCTGTATGAAGAGCAGCGGTGAGCGGCACTATCAAAACAACTGCGGCAAC 166
DB 222 AGCCCTTATGAGAGAGCAGCGGTGAGCGGCACTATCAAAACAACTGCGGCAAC 281
QY 167 GCTACGAGTCTTGTGAGAGCGCTGCGGCAAGGCACTTACGGAAGTGAAGAGCAAG 226
DB 282 GCTACGAGTCTTGTGAGAGCGCTGCGGCAAGGCACTTACGGAAGTGAAGAGCAAG 341

XX New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.

XX Claim 2, SEQ ID NO 639, 918bp, English.

XX This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytosolic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is a cDNA sequence which encodes a human PRO protein of the
CC invention.

XX Seq Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

XX Query Match 71.6%; Score 1355.4; DB 10; Length 3443;

XX Best Local Similarity 84.0%; Pred. No. 0; Mismatches 276; Indels 21; Gaps 2;

XX Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CCTCCGCGCTGCTCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCGCTTA 106
DB 162 CCACTCCCTCGCGCGAGAGTACCGCGCGCTGCGGAGCGGCTCATCAAGTCGCTTA 221
QY 107 AACCTGATGAGAGAGAGCGGCTGAGCGGCAACATCAACAACACTGCGGCAAC 166
DB 222 AGCCCTTAATGAGAGAGAGCGGCTGAGCGGCAACATCAACAACACTGCGGCAAC 281
QY 167 GCTACGAGTCTCTGAGAGCGCTGCGGCAAGGCACTTACGGGAGTGAAGAGCGAGAG 226
DB 282 GCTACGAGTCTCTGAGAGCGCTGCGGCAAGGCACTTACGGGAGTGAAGAGCGAGAG 341
QY 227 AGAGCTGCGGCGCTGCTGCTGCGCATCAAGTCATCAAGAGAGAGAGAGAGAGAG 286
DB 342 AGAGCTGCGGCGCTGCTGCTGCGCATCAAGTCATCAAGAGAGAGAGAGAGAGAG 401
QY 287 AGAGCTGCTGAG 346
DB 402 AAGATCTGATGCAATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
QY 347 TCATGCTGATGCAATGAG 406
DB 462 TCATGCTGATGCAATGAG 521
QY 407 CCAAGCGAGGCGATCTGATGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 466
DB 522 CCAAGCGAGGCGATCTGATGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 581
QY 467 CCAAGCGATTTTCTGCAAGAGTGTCTGCTGCACTACTGCAACCAAGAGAGAGAGAG 526
DB 582 CTAGGCAATTTCTTCCGAGAGATCTCTGCTGCTGCACTACTGCAACCAAGAGAGAG 641
QY 527 TTGACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586
DB 642 TTGACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 701
QY 587 CTGACTTTGAGCTCTCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646

DB 702 CTGACTTTGAGCTCTCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761
QY 647 GCCCTCTTCAAGCGCTGCGCTGAGATAGTCAACGAGAGAGAGAGAGAGAGAGAGAGAG 706
DB 762 GCCCTCTTCAAGCGCTGCGCTGAGATAGTCAACGAGAGAGAGAGAGAGAGAGAGAGAG 821
QY 707 ACAAGTGTCTCTGAG 766
DB 822 ACAAGTGTCTCTGAG 881
QY 767 GCGAGATCTTAAACACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
DB 882 GCGATGACCTTAAAGTCTTGTAGTAAACAGTACAGCAACGAGAGAGAGAGAGAGAG 941
QY 827 AGCGTCCAGATGCTGAG 886
DB 942 AACCTTGTATGCTGCTGAG 1001
QY 887 CCAACTGAGAGAGATGATGAG 946
DB 1002 CCAACTGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1061
QY 947 GAG 1006
DB 1062 GAG 1121
QY 1007 TGGGAG 1066
DB 1122 TGGTGAAG 1181
QY 1067 GCTTCTTCAAG 1126
DB 1182 GCTTCTTCAAG 1241
QY 1127 ATTCTCTTAAAG 1186
DB 1242 ATTCTCTTAAAG 1301
QY 1187 CTGAGAGATCTCTTCTGCGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
DB 1302 CTGAGAGATCTCTTCTGCGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 1247 AGAAG 1306
DB 1362 AGAAG 1421
QY 1307 TGCCTGATCTCAAG 1366
DB 1422 TGCCTGATCTCAAG 1472
QY 1367 TTAAG 1426
DB 1473 TTAAG 1532
QY 1427 GGGAG 1486
DB 1533 GGGAG 1592
QY 1487 CAG 1546
DB 1593 CAG 1652
QY 1547 CCGGAG 1606
DB 1653 CCGGAG 1712
QY 1607 CTTCCCATCTGAG 1666
DB 1713 CACTTGCCTCTGAG 1772
QY 1667 TGTCTTCCAG 1726

Db 1773 TGTCTCTGAGTCTTGTGACAGCTGGACTTGTGTAAGGCTCCAGAGCCCCCACTGC 1832
QY 1727 GGGGCTGTGTGTCTGTGGAACAACCTGAGGGGCTTTGAGCAAGCTTCCTCAGAG----- 1780
Db 1833 GGGGCTGTGTGTGTGTGGAACAACCTCAAGGGGCTTGAAGAGCCCCCTCAGAGGGCCCTG 1892
QY 1781 -----GTCTGAAGCGATGGTGGCAGGAATCTTGGGGGATAGCTGCTTTCTCTGACAG 1834
Db 1893 GAAAGCTGCTGAGGGGCTGGGGGAGATCTTTGGGGGACAGCTGCTTTCTCTGACAG 1952
QY 1835 ACTGCCAAGAGGTGACTGCAAGCCTTACAGACAAAGCCTTAGGATCTGCTCAAGGCTCAGC 1893
Db 1953 ACTGCCAAGAGGTGACAGCAGCCTTACCGACAGGCACTGAGGGTCTGCTCAAGGCTCACC 2011

Search completed: May 15, 2005, 03:21:46
Job time : 989.386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 06:02:36 ; Search time 818.38 Seconds
(without alignments)
11298.520 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015
Perfect score: 1893
Sequence: 1 atgagatcggtgcctact.....gaatctgtcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2902	6	AR279568 Sequence
2	1893	100.0	2902	6	AR562226 Sequence
3	1888.2	99.7	2027	6	AX381046 Sequence
4	1872.2	98.9	2917	10	BC033302 Mus muscu
5	1859	98.2	2938	10	BC046833 Mus muscu
6	1686.6	89.1	2900	10	BC081899 Rattus no
7	1683.4	88.9	2026	6	AX380986 Sequence
8	1528.4	88.9	2929	6	AX380960 Sequence
9	1528.4	80.7	291762	2	AC125887 Rattus no
10	1360.2	71.9	1884	6	AX407118 Sequence
11	1360.2	71.9	3353	6	AX407116 Sequence
12	1357.6	71.7	3360	6	AX642961 Sequence
13	1357	71.7	1887	6	CQ714327 Sequence
14	1357	71.7	3395	6	CQ782778 Sequence
15	1357	71.7	3395	6	BD127309 Sequence
16	1357	71.7	3395	9	AK074830 Primer fo
17	1357	71.7	3402	9	BC017306 Homo sapi
18	1355.4	71.6	3443	6	AX086949 Sequence
19	1355.4	71.6	3443	6	AX780351 Sequence

20	1338.8	70.7	2291	6	AX056394 Sequence
21	1230	65.0	1833	6	AX803413 Sequence
22	1071.4	56.6	1186	6	AX380972 Sequence
23	943.4	49.8	1186	6	AX381029 Sequence
24	923.4	48.8	252978	2	AC126983 Rattus no
25	840.2	44.4	2616	6	AX430359 Sequence
26	840.2	44.4	2619	6	AX430468 Sequence
27	723	38.2	179414	9	AC098933 Homo sapi
28	719.8	34.4	6389	9	BX537564 Homo sapi
29	651.8	34.4	1274	5	BX950299 Gallus ga
30	648	34.2	707	6	AX381050 Sequence
31	648	34.2	712	6	AX380990 Sequence
32	644.2	34.0	661	6	AX381052 Sequence
33	613.6	32.4	741	6	AX380991 Sequence
34	592	31.3	734	6	AX381051 Sequence
35	583.2	30.8	588	6	AX381056 Sequence
36	557.4	29.4	660	6	AX380992 Sequence
37	533	28.2	578	6	AX381054 Sequence
38	515	27.2	548	6	AX381055 Sequence
39	510.6	27.0	579	6	AX380994 Sequence
40	506.4	26.8	1886	6	AX213158 Sequence
41	506.4	26.8	2884	6	AX213160 Sequence
42	506.4	26.8	6825	6	CQ715220 Sequence
43	506.4	26.8	6828	6	CQ776374 Sequence
44	506.4	26.8	6828	6	AX213156 Sequence
45	506.4	26.8	6828	9	AB011109 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR279568 2902 bp DNA linear PAT 10-Apr-2003
DEFINITION Sequence 4 from patent US 6514719.
ACCESSION AR279568
VERSION AR279568.1 GI:29714427
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2902)
AUTHORS Bird,T.A., Vireca,G.D., Martin,U. and Anderson,D.M.
TITLES Methods for identifying compounds that alter kinase activity
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;
FEATURES
Source 1. 2902
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1893; DB 6; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGATCGGTGCTTAATCTCCAGCGCCGAGCAGGCTCCCTGCGCTCCGCTGAGC 60
DB 123 ATGAGATCGGTGCTTAATCTCCAGCGCCGAGCAGGCTCCCTGCGCTCCGCTGAGC 182
QY 61 TCGGAGAGCGCGCGCGCTGCGGAGCGGCTTCATCAAGTCGCTTAACTTGATGAAG 120
DB 183 TCGGAGAGCGCGCGCGCTGCGGAGCGGCTTCATCAAGTCGCTTAACTTGATGAAG 242
QY 121 AAGCAGCGCGTGAAGGCGCACCCTCAACAACCACTTCGCGGCTTCAGGTTCTTG 180
DB 243 AAGCAGCGCGTGAAGGCGCACCCTCAACAACCACTTCGCGGCTTCAGGTTCTTG 302
QY 181 GAGACCGCTGGGCAAGGCGCCTTACGAGGAGTGAAGAGGCAAGAGCTCGGGCGT 240
DB 303 GAGACCGCTGGGCAAGGCGCCTTACGAGGAGTGAAGAGGCAAGAGCTCGGGCGT 362
QY 241 CTGGTGGCATCAAGTCCATCAGGAAGAACAATCAAAATGAGAGGAGATCTGTCGAC 300

Db 363 CTGTGGCCATCAATGCATCAGAAAAGCAAAATCAAAATGAGAGATCTGTGCAC 422
QY 301 ATACGAGAGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 360
Db 423 ATACGAGAGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 482
QY 361 GAAGTGTGAGAAATGACAGCAAGATTTGATGATGAGATGATGACAGAGAGAGAT 420
Db 483 GAAGTGTGAGAAATGACAGCAAGATTTGATGATGAGATGATGACAGAGAGAGAT 542
QY 421 CTGTATGATTTACATCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
Db 543 CTGTATGATTTACATCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 602
QY 481 CGACAGATGTGTCTGTCCCTGCACTAATGACCAAGAGAGAGAGAGAGAGAT 540
Db 603 CGACAGATGTGTCTGTCCCTGCACTAATGACCAAGAGAGAGAGAGAGAT 662
QY 541 AAGCTGGAAGAAATCTTCTAGATGACAGAGAGAGAGAGAGAGAGAGAT 600
Db 663 AAGCTGGAAGAAATCTTCTAGATGACAGAGAGAGAGAGAGAGAGAGAT 722
QY 601 TCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 660
Db 723 TCCAACTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 782
QY 661 TGGCCCTGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 720
Db 783 TGGCCCTGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 842
QY 721 GGGCTTCTCTGTATCATCTGTGATGACAGAGAGAGAGAGAGAGAGAT 780
Db 843 GGGCTTCTCTGTATCATCTGTGATGACAGAGAGAGAGAGAGAGAGAT 902
QY 781 ACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840
Db 903 ACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 962
QY 841 TGTGGCTGTATCTGTGATGACAGAGAGAGAGAGAGAGAGAGAT 900
Db 963 TGTGGCTGTATCTGTGATGACAGAGAGAGAGAGAGAGAGAGAT 1022
QY 901 GTAGCAGATCATTTGTGAGAGAGAGAGAGAGAGAGAGAGAT 960
Db 1023 GTAGCAGATCATTTGTGAGAGAGAGAGAGAGAGAGAGAT 1082
QY 961 CTGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
Db 1083 CTGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1142
QY 1021 CGTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080
Db 1143 CGTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1202
QY 1081 CAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1140
Db 1203 CAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1262
QY 1141 TCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1200
Db 1263 TCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1322
QY 1201 TCTGGCCCTGCAAGAGAGAGAGAGAGAGAGAGAT 1260
Db 1323 TCTGGCCCTGCAAGAGAGAGAGAGAGAGAGAGAT 1382
QY 1261 ACCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1320
Db 1383 ACCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1442
QY 1321 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
Db 1443 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1502

QY 1381 CAGCGTGAATCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1503 CAGCGTGAATCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1562
QY 1441 GCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1563 GCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
QY 1501 CTCCTCTCCACCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1560
Db 1623 CTCCTCTCCACCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1682
QY 1561 GAAAGCATTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1620
Db 1683 GAAAGCATTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1742
QY 1621 GCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1680
Db 1743 GCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1802
QY 1681 TTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1803 TTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1862
QY 1741 GTGAGCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
Db 1863 GTGAGCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1922
QY 1801 GAATCTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1860
Db 1923 GAATCTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1982
QY 1861 AGACAAGCCCTTGAATCTGTCTCAAAAGCTCAGC 1893
Db 1983 AGACAAGCCCTTGAATCTGTCTCAAAAGCTCAGC 2015

RESULT 2
AR562226 2902 bp DNA linear PAT 08-OCT-2004
LOCUS AR562226
DEFINITION Sequence 4 from patent US 6759223.
ACCESSION AR562226
VERSION AR562226.1 GI:53976143
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2902)
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.
TITLE Calcium/calmodulin-dependent kinase
JOURNAL Patent: US 6759223-A 4 06-JUL-2004;
FEATURES
source location/Qualifiers
1..2902
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1893; DB 6; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGATGAGTGGCTTACTTCAAGAGAGAGAGAGAGAGAGAT 60
Db 123 ATGAGATGAGTGGCTTACTTCAAGAGAGAGAGAGAGAGAT 182
QY 61 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
Db 183 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 242
QY 121 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 180
Db 243 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 302

Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, A.M., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22389257

JOURNAL

PUBMED

12477932

2 (bases 1 to 2917)

Strausberg, R.

Direct Submission

Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulès, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>

Series: IRAC Plate: 40 Row: d Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229181.

Location/Qualifiers

FEATURES

source

1..2917

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="MGC:30598 IMAGE:3663595"

/tissue_type="Mammary tumor metastasized to lung."

MMTV-LTR/Mmt1 Model. Expression driven by an MMTV-LTR enhancer."

/clone_1ib="NCI CGAP_Lu30"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1..2917

/gene="1200013B22R1k"

/note="synonym: Snark"

/db_xref="locusid:74137"

/db_xref="MGI:1921387"

81..1976

/gene="1200013B22R1k"

/codon_start=1

/product="1200013B22R1k protein"

/protein_id="AAH33302.1"

/db_xref="MGI:23271086"

/db_xref="locusid:74137"

/db_xref="MGI:1921387"

/translation="MESVALQRPSSQAPSASALASBARLADGLIKSPKPLMKQAV KRHHKHLNHRYPETLIGKTYGKVKARSSGRLLVAIKSKIKDQDLHHR

REIEMSLNHPHIIAHEVENSXIVYVMEASRGDYDYSERPLSERDARHF
ROIVSALHCHONGVIVHRDKEMLILLDNGNLTADFGSLNHYHKEIQTGCSL
VAPSLVARKPYVGEVIVDPSGSLVILLVHGMGTGGDHLKVLQISGAARPEK
PSDAGLILMLVNPTRATLEVDVSHMMVMVGYTGGEOALREGGSDSGA
SNADMRSSRPLLENGARVCFKOHVGGSTVGLRSHLSKSRKENDANOIO
GPBARTSSRPPKSLIKPKGILKXSSSSGEBVDPBETRPVPTGPVAVSL
PRGILKSRKÖBSGTSPEPSEBELDASVFGPPEVKSQASGLLHRKCI
LKNGFSTALEGTPSTFGSLDQASSHPAPRPSGAVSEDSILSSSEDL
PRLPETPRGCVSVNLRGLRBPSPSGKRWQBSLGSFSLTDCOVTAVRQAL
GICSKLS"

249..1001

/gene="1200013B22R1k"

/note="S_TKc Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"

/db_xref="CDI:cd00180"

ORIGIN

Query Match 98.9%; Score 1872.2; DB 10; Length 2917;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1880; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

1 ATGGAATCGGTGGCTTACTCCAGGCGCCGAGCAGGCTCCCGGCTCGGCTGGCC 60

81 ATGGAATCGGTGGCTTACTCCAGGCGCCGAGCAGGCTCCCGGCTCGGCTGGCC 140

61 TCGGAAGAGCGCCGCGCGCTGCGGAGCGGCTCATGAAGTGGCTAACTGTGATGAAG 120

141 TCGGAAGAGCGCCGCGCGCTGCGGAGCGGCTCATGAAGTGGCTAACTGTGATGAAG 200

121 AAGCAGCGCGTGAAGCGGCGCATCAACAACACCTGGCGCAACGCTGATGAGTTCTTG 180

201 AAGCAGCGCGTGAAGCGGCGCATCAACAACACCTGGCGCAACGCTGATGAGTTCTTG 260

181 GAGAGCGTGGCGAAGGCGACCTACGGGAGGTGAAGAAGCGAGAGAGTGGGCGT 240

261 GAGAGCGTGGCGAAGGCGACCTACGGGAGGTGAAGAAGCGAGAGAGTGGGCGT 320

241 CTGGTGGCATCAAGTCCATCAGGAAGAACAATATGAGATGAGAGATGCTGCTGAC 300

321 CTGGTGGCATCAAGTCCATCAGGAAGAACAATATGAGATGAGAGATGCTGCTGAC 380

301 ATACGAGAGAGATGAGATGATGCTTCACTCAACCAACCCCAATATGCTGCTGAT 360

381 ATACGAGAGAGATGAGATGATGCTTCACTCAACCAACCCCAATATGCTGCTGAT 440

361 GAAGTGTGAGATGAGATGAGATGATGCTTCACTCAACCAACCCCAATATGCTGCTGAT 420

441 GAAGTGTGAGATGAGATGAGATGATGCTTCACTCAACCAACCCCAATATGCTGCTGAT 500

421 CTGTATGATTAATCATGATGAGCGGCGCATGATGAGCGGAGCGAGATTTCTTC 480

501 CTGTATGATTAATCATGATGAGCGGCGCATGATGAGCGGAGCGAGATTTCTTC 560

481 CGACAGATCGTGTCTGCGCTGCACTAATGCGACAGAGCGGATGCTTCAACGAGATCTC 540

561 CGACAGATCGTGTCTGCGCTGCACTAATGCGACAGAGCGGATGCTTCAACGAGATCTC 620

541 AAGCTGAAAACATCTCTTGAATGCGATGAAAACATCAAGATTTGCTGACTTGGCTC 600

621 AAGCTGAAAACATCTCTTGAATGCGATGAAAACATCAAGATTTGCTGACTTGGCTC 680

601 TCGACCTGACCAAGAGCAAGTCTTCAAGAGTCTGAGAGCGCTCTCAAGCC 660

681 TCGACCTGACCAAGAGCAAGTCTTCAAGAGTCTGAGAGCGCTCTCAAGCC 740

661 TCGCCTGAGATGACAGGGAAGCCCTATGTTGGGCGCAAGGTGACAGTGTGCTCTG 720

Db 741 TCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGACAGCTGTCTCTG 800
 Qy 721 GGGGTCTCTCTGATCAATCCCTGGTGCATNGGACCAATGCCCTTTTGACGGGAGATCATATA 780
 Db 801 GGGGTCTCTCTGATCAATCCCTGGTGCATNGGACCAATGCCCTTTTGACGGGAGATCATATA 860
 Qy 781 ACACTGTGAGAGCAATCAAGTAAACGGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCC 840
 Db 861 ACACTGTGAGAGCAATCAAGTAAACGGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCC 920
 Qy 841 TGTGCTCTGATCCGGTGGCTGTTAATGTGAACCCCAACCGTGGGCTCACATGAGAGAT 900
 Db 921 TGTGCTCTGATCCGGTGGCTGTTAATGTGAACCCCAACCGTGGGCTCACATGAGAGAT 980
 Qy 901 GTAGCCAGATCATTTGGTGGCTGCACTGGGCTTAAACCAACCGAGTGGGGAACAGAAACC 960
 Db 981 GTAGCCAGATCATTTGGTGGCTGCACTGGGCTTAAACCAACCGAGTGGGGAACAGAAACC 1040
 Qy 961 CTGCTGAGAGGTGGGACCCCTAGTGGTGACTTTGGCCGGGCTCCANTGGGAGCTGGTGA 1020
 Db 1041 CTGCTGAGAGGTGGGACCCCTAGTGGTGACTTTGGCCGGGCTCCANTGGGAGCTGGTGA 1100
 Qy 1021 CGTGGCTCTCGGCCCCCTCTCGAGAGATGAGCCAGGTGTGACGCTTCTTCAAGCAG 1080
 Db 1101 CGTGGCTCTCGGCCCCCTCTCGAGAGATGAGCCAGGTGTGACGCTTCTTCAAGCAG 1160
 Qy 1081 CAGGTGGGAGGTGGGAGGATCACTGTAACCTGGGCTGAGCGGCAATTTCTTAAAGAG 1140
 Db 1161 CAGGTGGGAGGTGGGAGGATCACTGTAACCTGGGCTGAGCGGCAATTTCTTAAAGAG 1220
 Qy 1141 TCCCGAAGAGAGATGATCATGTGCTCAAAATCTGCAAGGTGACCCGGGTGAGATACCTCT 1200
 Db 1221 TCCCGAAGAGAGATGATCATGTGCTCAAAATCTGCAAGGTGACCCGGGTGAGATACCTCT 1280
 Qy 1201 TCTGCTCTGGCAAGAGAGGCTTAAAGCTTCCGAAAGCATTTCTCAAGAAAGTCTCTCT 1260
 Db 1281 TCTGCTCTGGCAAGAGAGGCTTAAAGCTTCCGAAAGCATTTCTCAAGAAAGTCTCTCT 1340
 Qy 1261 ACTGCTGAGGAGAGGTGATCAAGAGAGAGCCCTGAGAGAACTGACCGGTGCTGATCTCA 1320
 Db 1341 ACTGCTGAGGAGAGGTGATCAAGAGAGAGCCCTGAGAGAACTGACCGGTGCTGATCTCA 1400
 Qy 1321 GGGCAGCCTGTCTCTGTGATCCCTGTCTCCAGAGAGGAGCATCTTAAAGATCTCGA 1380
 Db 1401 GGGCAGCCTGTCTCTGTGATCCCTGTCTCCAGAGAGGAGCATCTTAAAGATCTCGA 1460
 Qy 1381 CAGCGTGAATCTGTGATCTCTCTCTCAAGAGCCGAGTCTGGGGAATCTTAAAGAC 1440
 Db 1461 CAGCGTGAATCTGTGATCTCTCTCTCAAGAGCCGAGTCTGGGGAATCTTAAAGAC 1520
 Qy 1441 GGCAGTGAATCTGTGATCTGTGAGTGGGAGCCCGGTGAGAGAGATCTTCAAGAGG 1500
 Db 1521 GGCAGTGAATCTGTGATCTGTGAGTGGGAGCCCGGTGAGAGAGATCTTCAAGAGG 1580
 Qy 1501 CTCTCTCTCAACCGAGAGGAGCATTTCAAACTCAATGAGGAGTCTCCGCAAGGCTTA 1560
 Db 1581 CTCTCTCTCAACCGAGAGGAGCATTTCAAACTCAATGAGGAGTCTCCGCAAGGCTTA 1640
 Qy 1561 GAAGGACATACCTCTGAGCATTTGGGCTCTCTGAGCAACCTGAGCTCTCTCCATCTGCA 1620
 Db 1641 GAAGGACATACCTCTGAGCATTTGGGCTCTCTGAGCAACCTGAGCTCTCTCCATCTGCA 1700
 Qy 1621 GCCCGGAGAGAGCCCTCTCAAGGAGGCTGTGAGTGAAGAGACATCTGTCTCCAGTCC 1680
 Db 1701 GCCCGGAGAGAGCCCTCTCAAGGAGGCTGTGAGTGAAGAGACATCTGTCTCCAGTCC 1760
 Qy 1681 TTTGACCAATTTGAGCTTGTGCTGAAGCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1740
 Db 1761 TTTGACCAATTTGAGCTTGTGCTGAAGCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1820
 Qy 1741 GTGACCAACTGAGGGGCTTGAAGAGCTCTCTCAAGAGGTCTGAAGGATGTGGCAG 1800
 Db 1821 GTGACCAACTGAGGGGCTTGAAGAGCTCTCTCTCAAGAGGTCTGAAGGATGTGGCAG 1880

Qy 1801 GAATCTTGGGGGATAGCTGCTTTTCTTGAAGAGATGCCAGAGGTGACTGACCTTAC 1860
 Db 1881 GAATCTTGGGGGATAGCTGCTTTTCTTGAAGAGATGCCAGAGGTGACTGACCTTAC 1940
 Qy 1861 AGACGAGCCCTAGGATCTGCTCAAGCTCAGC 1893
 Db 1941 AGACGAGCCCTAGGATCTGCTCAAGCTCAGC 1973

RESULT 5
 BC046833
 LOCUS
 DEFINITION
 Mus musculus RIKEN cDNA 1200013B22 gene, mRNA (CDNA clone MGC:61244
 IMAGE:6816569), complete cds.
 ACCESSION
 BC046833
 VERSION
 BC046833.1 GI:29145027
 KEYWORDS
 MGC.
 SOURCE
 Mus musculus (house mouse)

REFERENCE
 1 (bases 1 to 2938)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Burow, K.H., Scheaffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skaleja, V., Small, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 2938)
 Strausberg, R.
 Direct Submission
 Submitted (13-FEB-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
 Fishler, K., Kappel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
 Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
 Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 27229181.

FEATURES
source
Location/Qualifiers
1..2938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="WGC:61244 IMAGE:6816569"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_1ib="NIH BMAP_FW0"
/lab_host="DH10B"
/note="Vector: pYX-Asc"
1..2938
/gene="1200013B22Rik"
/note="synonym: Shark"
/db_xref="LocusID:74137"
/db_xref="WGI:1921387"
81..2000
/gene="1200013B22Rik"
/product="1200013B22Rik protein"
/codon_start=1
/protein_id="AAH4683.1"
/db_xref="GI:29145028"
/db_xref="LocusID:74137"
/db_xref="WGI:1921387"
/translation="MESVALILQRPQSPASASALASBPALDGLIKSPKPLMKQAV
KHHKHNLRHYELETGKTYGKVKARSSGRVAKISIRKDKIDEDDLHIR
REIIMSSLNHPIITAIHEVGRSLVTVPENSKYIVWEVASRGLDYIISERPLS
REDARPRQVYSAHYCHONGIYHRDLLENILIDANGNITIDANGLSNLYHKRFL
QTPGSPLYASPEIYNGKPYVPEVDNSELGVLYITLVHGTMPFDQDKHTLVKQISN
GAYRBPPEKSDACGLIRMLWNPYRATLEVDVASHWVNWQYTGVEGQELAREGH
PSGDFRASMADMLRRSRPLLENGAKVCSFKOHVGGGSLVPGIERHSLKSKKE
NDMAONLQDPAEDTSSRPGKSLKLPKILKKKSTSGEVOEDPELRPYDPTFG
PPAYSLIPRKGILKKSRORSEGYVSPSPSGELIDSDVFSGDPIYKOSPOAG
LLHKGITIKANGKESRTLEGTPTSTSDQLASHPAARPSRSGAVSDSTLSS
ESFDQDPERLPETPLRGCVSDNLRGLPEPSPBGLKRWQESLSDSCPSLTDCEV
TAYRQALGICSKLS"
249..1025
/gene="1200013B22Rik"
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"
ORIGIN
Query Match 98.2%; Score 1859; DB 10; Length 2938;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 24; Gaps 1;
QY 1 ATGAGTCCGTGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 60
DB 81 ATGAGTCCGTGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 140
QY 61 TCGAGAGCGCCGCTGGCGGAGCGGCTCATCAAGTCGCTAAACCTTCGAGGAAG 120
DB 141 TCGAGAGCGCCGCTGGCGGAGCGGCTCATCAAGTCGCTAAACCTTCGAGGAAG 200
QY 121 AAGCAGGCGGCTGAAGCGGACCATCAACAACAACCTGCGGACCGGCTACGATCCCG 180
DB 201 AAGCAGGCGGCTGAAGCGGACCATCAACAACAACCTGCGGACCGGCTACGATCCCG 260
QY 181 GAGACGCTGGCAAGGCACTTACGGGAAGTGAAGAAGCAAGAGAGCTCGGGCGT 240
DB 261 GAGACGCTGGCAAGGCACTTACGGGAAGTGAAGAAGCAAGAGAGCTCGGGCGT 320
QY 241 CTGATGGCCATCAAGTCCATCAGGAAGAATAAATCAAGATAGCAAGATCTGCTGCAC 300
DB 321 CTGATGGCCATCAAGTCCATCAGGAAGAATAAATCAAGATAGCAAGATCTGCTGCAC 380

QY 301 ATACGAGGAGATTTGAGTATCATGTCTTACCTCAACCAACCCCACTATTCGATCCAT 360
DB 381 ATACGAGGAGATTTGAGTATCATGTCTTACCTCAACCAACCCCACTATTCGATCCAT 440
QY 361 GAA-----GTTTGAAGATAGAGCAAGATTGATTTGTC 396
DB 441 GAAATGGGCAATCAACGCTTAATGACTGTGTTGAGATAGAGCAAGATTGATTTGTC 500
QY 397 ATGAGTATGACGAGCCGAGCCGATCTGTATGATTAATCATCATGTAAGGCGCAAGCTGAGT 456
DB 501 ATGAGTATGACGAGCCGAGCCGATCTGTATGATTAATCATCATGTAAGGCGCAAGCTGAGT 560
QY 457 GAGCGGAGCCGAGGCAATTTCTTCCGACGATGTGTCTGCTTCGCTGACATCTGACCAAG 516
DB 561 GAGCGGAGCCGAGGCAATTTCTTCCGACGATGTGTCTGCTTCGCTGACATCTGACCAAG 620
QY 517 AACGGATCGTTCACCGGATCTCAAGCTGGAAGAAACATCTTCTGATGCCAATGGAAAC 576
DB 621 AACGGATCGTTCACCGGATCTCAAGCTGGAAGAAACATCTTCTGATGCCAATGGAAAC 680
QY 577 ATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCCAGAG 636
DB 681 ATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCCAGAG 740
QY 637 TTCTGTGGAGCCCTCTCTACGCTTGGCTGAGATAGTCAACGGGAAGCCTTATGTGGC 696
DB 741 TTCTGTGGAGCCCTCTCTACGCTTGGCTGAGATAGTCAACGGGAAGCCTTATGTGGC 800
QY 697 CCAGAGTGGACAGCTGTCTGTGGCCGTTCTCTGTATCATCTGATGATGGCAACAG 756
DB 801 CCAGAGTGGACAGCTGTCTGTGGCCGTTCTCTGTATCATCTGATGATGGCAACAG 860
QY 757 CCCTTGAAGGCGAGGATCATTAACAACATGTGAAGCAATCATGTAACGGGCTTACCGT 816
DB 861 CCCTTGAAGGCGAGGATCATTAACAACATGTGAAGCAATCATGTAACGGGCTTACCGT 920
QY 817 GAGCGCCCAAGCCGTCGATGCTGTGGCTGATTCGGTGGCTGTAAATGTGAACCC 876
DB 921 GAGCGCCCAAGCCGTCGATGCTGTGGCTGATTCGGTGGCTGTAAATGTGAACCC 980
QY 877 ACCCGTGGGCACTGTGAGATGTAGCCAGTCAATGGTGGTCAACTGGGTTACCC 936
DB 981 ACCCGTGGGCACTGTGAGATGTAGCCAGTCAATGGTGGTCAACTGGGTTACCC 1040
QY 937 ACCGAGTGGGGAACAGGAGCCCTGCGTGAAGGCGGACCTTATGTGTGACTTTGGC 996
DB 1041 ACCGAGTGGGGAACAGGAGCCCTGCGTGAAGGCGGACCTTATGTGTGACTTTGGC 1100
QY 997 CGGCGCTCCATGCGGACCTGTAAAGTCTGCTCTCGCGCCCTCTCTGAGATGAGCC 1056
DB 1101 CGGCGCTCCATGCGGACCTGTAAAGTCTGCTCTCGCGCCCTCTCTGAGATGAGCC 1160
QY 1057 AAGGTGTGAGCTTTCTTCAAGCAAGCTGCGGAGGTGAAGCACTGTACTGGGCTG 1116
DB 1161 AAGGTGTGAGCTTTCTTCAAGCAAGCTGCGGAGGTGAAGCACTGTACTGGGCTG 1220
QY 1117 GAGCGGCAATTTCTTAAAGAGTCCGAAAGGAGAAATGAGATGCTCAAAATCTGCA 1176
DB 1221 GAGCGGCAATTTCTTAAAGAGTCCGAAAGGAGAAATGAGATGCTCAAAATCTGCA 1280
QY 1177 GGTGACCGGCTGAGGATCTCTTCTCGCTGCGGCAAGAGAGCTTAAAGCTTCGAA 1236
DB 1281 GGTGACCGGCTGAGGATCTCTTCTCGCTGCGGCAAGAGAGCTTAAAGCTTCGAA 1340
QY 1237 GGCATTCTCAAGAAAAAGTCTTCTTCACTGTCAAGGAGGTACAGAGAGCTTCAGAA 1296
DB 1341 GGCATTCTCAAGAAAAAGTCTTCTTCACTGTCAAGGAGGTACAGAGAGCTTCAGAA 1400
QY 1297 CTCAAGCGGAGTCTGATCTCAAGGAGCGGCTGCTGCTGATACCTGCTCCCAAG 1356
DB 1401 CTCAAGCGGAGTCTGATCTCAAGGAGCGGCTGCTGCTGATACCTGCTCCCAAG 1460
QY 1357 AAAGGATCTTAAAGATCTGAGACAGTGAATCTGTACTTCTCTCAAGAGCC 1416

Db	1461	AAAGGATCCTTAAGAGTCTCGACAGCGTAATCTGTACTACTCTCTCCAGAGCCC	1520
Qy	1417	AGGAGTCTGGGAACTCTTAAGACGCCAGTATGTGTTGTAGTGGGACCCGTGGAG	1476
Db	1521	AGGAGTCTGGGAACTCTTAAGACGCCAGTATGTGTTGTAGTGGGACCCGTGGAG	1580
Qy	1477	CAGAGTCTCAGAGGCTTCAGGGCTCTCTCCACCGGAAAGGCAATCTCAATCAAT	1536
Db	1581	CAGAGTCTCAGAGGCTTCAGGGCTCTCTCCACCGGAAAGGCAATCTCAATCAAT	1640
Qy	1537	GCGAGTCTCCGCAAGCCTTAAGAGGCACTACCCCTAGACCTTTGGCTCCCTGGAC	1596
Db	1641	GCGAGTCTCCGCAAGCCTTAAGAGGCACTACCCCTAGACCTTTGGCTCCCTGGAC	1700
Qy	1597	CAGCTGGCTCTCTCCCTCACTCTGCAAGCCCGGCCAGCCCTCTAGAGGGCTGTAGTAG	1656
Db	1701	CAGCTGGCTCTCTCCCTCACTCTGCAAGCCCGGCCAGCCCTCTAGAGGGCTGTAGTAG	1760
Qy	1657	GACAGATCTCTCTCTCCAGTCTCTTGAACCAATGGAATGCTGTAAGCTTCCCGAA	1716
Db	1761	GACAGATCTCTCTCTCCAGTCTCTTGAACCAATGGAATGCTGTAAGCTTCCCGAA	1820
Qy	1717	ACCCCACTGAGGGGCTGTGTCTGTGACCACTGAGGGGGCTTGAAGACCTCTCTCA	1776
Db	1821	ACCCCACTGAGGGGCTGTGTCTGTGACCACTGAGGGGGCTTGAAGACCTCTCTCA	1880
Qy	1777	GAAAGTCTGAAGAGTGTGTGCGAGGAATCTTTGGGGGATAGCTGCTTTCTTGAACAC	1836
Db	1881	GAAAGTCTGAAGAGTGTGTGCGAGGAATCTTTGGGGGATAGCTGCTTTCTTGAACAC	1940
Qy	1837	TGCGAAGAGGTGACTGACGCTCAGACGACCAAGCCCTAGGAATCTGCTCAAGCTCAGC	1893
Db	1941	TGCGAAGAGGTGACTGACGCTCAGACGACCAAGCCCTAGGAATCTGCTCAAGCTCAGC	1997
RESULT 6			
BC081899		2900 bp mRNA linear ROD 02-SEP-2004	
LOCUS	BC081899	Rattus norvegicus cDNA clone MGC:93817 IMAGE:7110518, complete cds.	
DEFINITION	BC081899	GI:51859485	
ACCESSION	BC081899.1		
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 (bases 1 to 2900)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klausner, R.D., Collins, P.S., Wagner, L., Shennan, C.M., Schlier, G.D.,			
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, I., Wang, J., Hsieh, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,			
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
Carninci, P., Frange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,			
Abrahamson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,			
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W.,			
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Fahy, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S.,			
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,			
Schuerch, A., Schein, J.B., Jones, S.J. and Marra, M.A.			
human and mouse cdna sequences			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cdna sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
1 (bases 1 to 2900)			
Director MGC Project.			

TITLE	Direct Submission	
JOURNAL	Submitted (01-SEP-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov	
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Howard Jacobs cDNA Library Preparation: Express Genomics cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-sngc.stanford.edu Contact: (Dickson, Mark) mdedpax11.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAP Plate: 182 Row: a Column: 17 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein. Location/Qualifiers 1..2900 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /clone="MGC:93817 IMAGE:7110518" /tissue_type="Kidney, rat (Brown Norway)" /clone_1b="NIH_MGC_235" /lab_host="DH10B" /note="Vector: pEXpress1" 73..1965 /codon_start=1 /product="Unknown (protein for MGC:93817)" /protein_id="AAH81899.1" /db_xref="GI:51859486" /translation="MESVALHRGNLAPASALATESARPLADRLIKSPKPMKQAV KHHHKNLRRHREPLETGGTGVKVRSSGSLVAKIRKDKIDBDLHHR REILMSLNHPHIIAIIHEHPENSSKIVIVMAYARGDIYDYSRPRINERDAHFP ROIVSAHCHONGIVHRDLKENTILILASIKYIADPRLSYHKGRKLTQFCSPY IASPEIVANKPYVGPVDSWSLGLVLYLIVHGTMPDQDHTLVKQISSGAIVRECK PSADAGLRIRMLMNVPIRATLLEDAVASHVMVWGYSIRIGEBALREGBSGSRA SHADWLRSSRPLLENGAKVCSFPKQHYVPGSGSTGRLEROSLKSRENMAGTLO NPDADTSRPPKNSLKLKPKGILKKASPSGVEGPOELPVNSTPQOPVAPYPII PRKGLIKSRRORESGYVSPSPESGEILNADVDVSGDPMWOKSPQASGRHARGIIL KNGFRSRALSGTAPSTFGSLDQLASHPRTARARPSGAVESDILSSSESDQDLP ERUPETPLKGCVSVDNLRLKLEQPPSEGLKRWQESLGDSFSLTDCQEVTAAYRQALG ICSKLS"	
ORIGIN		
Query Match	89.1%; Score 1686.6; DB 10; Length 2900;	
Best Local Similarity	93.6%; Pred. No. 0;	
Matches 1771; Conservative	0; Mismatches 119; Indels 3; Gaps 1;	
Qy	1 ATGAGTCTGGTGGCTTAAGAGTCTCGACGCGCCGAGCAAGCTCCCTGGGCTCCGCGCTGGCC	60
Db	73 ATGAGTCTGGTGGCTTAAGAGTCTCGACGCGCGGAACTGGCTCTCCCTGGGCTCCGCGCTGGCC	132
Qy	61 TCGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTGGCTTAACCTGTATGAAG	120
Db	133 ACGGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTGGCCAAACCTGTATGAAG	192
Qy	121 AAGCAGCGCGTAAACCGGACCATCAACAACAACCTGGGACCGCTCAAGATTCCTG	180
Db	193 AAGCAGCGCGTAAACCGGACCATCAACAACAACCTGAGGACCGCTCAAGATTCCTG	252
Qy	181 GAGAGCTGGGCAAGGCGCACTACGCGAAGGTGAAGAGGCAAGAGACTCGGGCGCT	240
Db	253 GAGAGCTGGGCAAGGCGCACTACGCGAAGGTGAAGAGGCAAGAGACTCGGGACCG	312

OY	241	CTGGGGGCATCAAGTCCATCAGGAAAAACAAATCAAGATGAGACAGATCTGTGCAC	300
Db	313	CTGGTGGCCATCAAGTCTATATACGAGGAAGACAAATCAAGATGAGACAGATCTGTGCAC	372
OY	301	ATAAGAGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATATTGCCATCCAT	360
Db	373	ATAAGAGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATATTGCCATCCAT	432
OY	361	GAAGTCTTTGAGATAGACAGACAGATTGTGATTGTCAATGAGATATGCCAGCCGAGCGAT	420
Db	433	GAAGTCTTTGAGAACAGACAGCAAGATTGTGATTGTCAATGAGATATGCCAGCCGAGCGAT	492
OY	421	CTGATATGATTACATCAGTGAAGCGGCAAGCGGTGAGTGAAGCGGGAACCGACGGCATTTCTTC	480
Db	493	CTGATCAGTATCATCAGTGAAGCGGCAAGCGGTGAGTGAAGCGGGAACCGACGGCATTTCTTC	552
OY	481	CGACAGATCTGTCTGCGCCCTGCATCACTACGCCACAGAACGGGATCGTTCAACCGAGATCTC	540
Db	553	CGACAGATCTGTCTGCGCCCTGCATCACTACGCCACAGAACGGGATCGTTCAACCGGACCTC	612
OY	541	AAGCTGAAAAATCTCTTCTAGATGCCAATGAAAAATCAAGATTGTCTGACTTTGGCCTC	600
Db	613	AAGCTGAGAAACATCTCTTAGATGCCAGTGGCAACATCAAGATTGTCTGATTGGCCTC	672
OY	601	TCCAACCTGTATCAACAAGGCAAGTTCTCCAGAGCTTCTGTGGGAGGCCCTCTCTAGCGCC	660
Db	673	TCCAACCTGTATCAACAAGGCAAGTTCTCTCAGAGCTTCTGTGGAGGCCCTCTCTTAGTGC	732
OY	661	TCGCTGAGATAGTACCAAGGGAAGCCATATGAGGAGCCAGAGAGTGAACAGCTGGTCTGTG	720
Db	733	TCACCTGAGATCTGTACAGGGAAGCCCTATGTGGAGCCAGAGAGTGAACAGCTGGTCTGTG	792
OY	721	GAGCTTCTCTGTATCATCTGTGATGATGGACAATGCCCTTTTGAACGGGACAGATCATATA	780
Db	793	GAGCTTCTCTGTATCATCTGTGATGATGGACAATGCCCTTTTGAACGGGACAGATCATATA	852
OY	781	ACACTGTGTGAAGCAATATGATGAAGGGGCTTACCGGTGAAGCCGCCCAAGCCGTCCGATGCC	840
Db	853	ACCTGTGTGAACAATATGATGAAGGGGCTTACCGAGAGCCGTGCAACCCCTCTGATATGCC	912
OY	841	TGTGAGCCGTATCCGGGTGGCTGTATATGTGTGAACCCACCGGTGGGGACCACTGGAAGAT	900
Db	913	TGTGAGCCGTATCCGGGTGGCTGTATATGTGTGAATCCATCCGTGGGACCACTGTGAAGAT	972
OY	901	GTAACAGATCATTTGGTGGTCAACTGGGGTTTACACACCGAGTCCGGGAACAGAAAGCC	960
Db	973	GTAACAGATCATTTGGTGGTCAACTGGGGTTTACACACCGAATTTGGGGAACAGAAAGCT	1032
OY	961	CTGCGTGAAGGTTGGGCACTTAGTGTGACTTTGGCCGGGCTTCATGTGGCGGACTGTGTTA	1020
Db	1033	CTGCGAGAGGTTGGGCACTTAGTGTGACTCTGGCCGGGCTTCATGTGGCGGACTGTGTTA	1092
OY	1021	CGTGGCTCTTGGCGCCCTCTCTGTGAATAAGGATCGAAGGTGTGACGTTCTTCAAGAG	1080
Db	1093	CGTGGCTCTTGGCGCCCTCTCTGTGAAGATGGAAGCAAGTGTGACGTTCTTCAAGAG	1152
OY	1081	CACGTGCGGAGGTGAAGACATGTACTCTGGGCTGGAAGCGGCAACATTTCTTAAAGAG	1140
Db	1153	CATGTGCGGAGGTGTGAAGACACGGGACCGGGGCTGGAAGCGGCAACATTTCTTAAAGAG	1212
OY	1141	TCGCCAAGAGAAATGATCATGTGCTCAAAATCTGCAAGGTGAACCGGCTGAGGATACCTT	1200
Db	1213	TCGCCAAGAGAAATGATCATGTGCTCAAGCTCTGCGAATGACCAAGCTGGAAGATACCTTC	1272
OY	1201	TCTGGCCCTTGGCAAGAGGACTTAAGCTTCGAAAGGCAATTCACAAATAATCCCTCT	1260
Db	1273	TCTGGCCCTTGGCAAGAGGACTTAAGCTTTCGAAAGGTAATCTCAAAAAAAGGCTCT	1332
OY	1261	ACCTGTCAAGGAGAGGTACAGAGAGACCTTAGAAGACTCAAGCGGTGCTGATATCTTCA	1320
Db	1333	CCCTATTCGGGGAGAGTACAGAGAGGCTCTAGAGAACTCAAGCAAGTGTCCAAATACCCCA	1392
OY	1321	GGGCAAGCTGTTCCTGTGTATCCCTGTCCCAAGAAAGGCAATCTTAAAGATCTTCA	1380

[illegible]

Db 135 ACCGAGAGCGCCGCGCTGGCGGACCGGCTCATGAATGCGCCAAACCTCTGATGAG 134
QY 121 AAGCAGCGCGTGAAGCGGCACTCAACAAACAACCTGGCGACCGCTACGAGTTCTG 180
Db 195 AAGCAGCGCGTGAAGCGGCACTCAACAAACAACCTGGCGACCGCTACGAGTTCTG 254
QY 181 GAGACGCTGGGCAAGGCACTACGGGAGGTGAAGAGGCAAGAGAGCTGGGGCGT 240
Db 255 GAGACGCTGGGCAAGGCACTACGGGAGGTGAAGAGGCAAGAGAGCTGGGGCGC 314
QY 241 CTGGTGCCCATCAAGTTCATCAAGAAAGCAAAATCAAGATGAGCAGATCTGCTGAC 300
Db 315 CTGGTGCCCATCAAGTTCATCAAGAAAGCAAAATCAAGATGAGCAGATCTGCTGAC 374
QY 301 ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 360
Db 375 ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 434
QY 361 GAAGTGTGAGATGAGCAGCAAGATTGTGATTGTCAATGAGTATGCGACCGAGGCGAT 420
Db 435 GAAGTGTGAGATGAGCAGCAAGATTGTGATTGTCAATGAGTATGCGACCGAGGCGAT 494
QY 421 CTGTATGATTATCATAGTGAAGCGGCAAGCTGAGTGAAGCGGCAAGCAGCATTCTTC 480
Db 495 CTGTATGATTATCATAGTGAAGCGGCAAGCTGAGTGAAGCGGCAAGCAGCATTCTTC 554
QY 481 CGACATATGCTGTCTGCGCTGCACTACTGCAACCAAGCGGATCTTCAACCGAGTCTC 540
Db 555 CGACATATGCTGTCTGCGCTGCACTACTGCAACCAAGCGGATCTTCAACCGAGTCTC 614
QY 541 AAGCTGAAAACATCTCTTGAATGCAATGAAACATCAAGATTGCTGACTTTGGCTTC 600
Db 615 AAGCTGAAAACATCTCTTGAATGCAATGAAACATCAAGATTGCTGACTTTGGCTTC 674
QY 601 TCCAACTGTATCCAAAGGCAAGTCTTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 660
Db 675 TCCAACTGTATCCAAAGGCAAGTCTTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 734
QY 661 TCGCCCTGAATATCAACCGGAAAGCCCTATGTGGGCGCCAGAGGTGAACAGTGTCTCTG 720
Db 735 TCACTGAGATCTGCAACCGGAAAGCCCTATGTGGGCGCCAGAGGTGAACAGTGTCTCTG 794
QY 721 GGGGTTCTCTGTAATCTCTGATGCAATGCAACATGCGCTTGAACCGGAGGATCAATAA 780
Db 795 GGGGTTCTCTGTAATCTCTGATGCAATGCAACATGCGCTTGAACCGGAGGATCAATAA 854
QY 781 ACACTGTGAAGCAAAATCAATGCAACCGGAGCTTACCGTGAAGCGGCAAGCGCTCGATGCC 840
Db 855 ACCCTGTGAACAAATCAATGCAACCGGAGCTTACCGAGAGCGCTGATGCC 914
QY 841 TGTGGCTGTATCCGTTGCTGTTAATGTGAACCCCAACCGTGGGCGACATGAGAGAT 900
Db 915 TGTGGCTGTATCCGTTGCTGTTAATGTGAATCCATCGTGGGCGACATGAGAGAT 974
QY 901 GTAGCAGATCATTTGGTGTCAACTGGGGTTTACCAACCGGATCGGGGAAACGAAAGCC 960
Db 975 GTAGCAGATCATTTGGTGTCAACTGGGGTTTACCAACCGGATCGGGGAAACGAAAGCT 1034
QY 961 CTGCGTGAAGGTGGGCAACCTAGTGTGAATTGGCGCGGCGCTCAATGGCGAGTGTGTA 1020
Db 1035 CTGCGGAGAGGTGGGCAACCTAGCGGTGAATCTGGCGGCGCTCTATGGCGAGCTGTGTA 1094
QY 1021 CGTGGCTCTCTCGCGCCCTCTCTGAGAAATGAGCCAAAGTGTGACGCTTCTTCAAGCAG 1080
Db 1095 CGTGGCTCTCTCGCGCCCTCTCTGAGAAATGAGCCAAAGTGTGAGCTTCTTCAAGCAG 1154
QY 1081 CAGGTGCGGGAGGTGGAGCACTGTACTGTGGCTGAGAGGCAACTTCTTTAAGAG 1140
Db 1155 CAGGTGCGGGAGGTGGAGCACTGTACTGTGGCTGAGAGGCAACTTCTTTAAGAG 1214
QY 1141 TCCCGAAAGAGATGATCATGTGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTCT 1200

Db 1215 TCCCGAAAGAGATGATCATGTGCTCAGACTCTGCAAGATGACCAGTTGAATACCTTCC 1274
QY 1201 TCTGCGCTCTGGCAAGGCAAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAGTCTCTC 1260
Db 1275 TCTGCGCTCTGGCAAGGCAAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAGTCTCTC 1334
QY 1261 ACCTGTCAGGGAGGTGAAGAGAGCCCTCAGAGAACTAGAGCCGGTGCCTGATCTCCA 1320
Db 1335 CCTCATCGGGAGGTGAAGAGAGCCCTCAGAGAACTAGAGAACTAGAGAACTAGATACCCA 1394
QY 1321 GGGCAGCTGTCTCTGCTGATCTCTGCTCCCAAGGAAAGGCAATCTTAAAGATCTCCA 1380
Db 1395 GGGCAGCTGTCTCTGCTGATCTCTGCTCCCAAGGAAAGGCAATCTTAAAGATCTCCA 1454
QY 1381 CAGCTGAATCTGTGTTACTACTCTCTTCCAGAGCCCAAGAGAGTCTGGGAACTCTTGAAC 1440
Db 1455 CAGCTGAATCTGTGTTACTACTCTCTTCCAGAGCCCAAGAGAGTCTGGGAACTCTTGAAC 1514
QY 1441 GCCAGTATGTGTTTGAAGTGGGAGCCCGTGGAGCAGAAATCTCAGAGGCTTCAAGG 1500
Db 1515 GCGAGTATGTGTTTGAAGTGGGAGCCCGTGGAGCAGAAATCTCAGAGGCTTCAAGG 1574
QY 1501 CTCCTCTCCACCGCAAGGCAATCTCAAACTCAATGCGAAGTCTCCGCAAGCCTTA 1560
Db 1575 ---CGCTCATCTGCAAGGCAATCTCAAACTCAATGCGAAGTCTCCGCAAGCCTTA 1631
QY 1561 GAAAGCACTACCTCTAGACCTTTGGCTCTCTGAGAACATGCGCTCTCCATCTCTGA 1620
Db 1632 GAAAGCACTACCTCTAGACCTTTGGCTCTCTGAGAACATGCGCTCTCCATCTCTGA 1691
QY 1621 GCCCGGCGAGCGCCCTCAGGGGCTGAGTGAAGAGCAGAACTCTGCTCCGAGTCC 1680
Db 1692 GCCCGGCGAGCGCCCTCAGGGGCTGAGTGAAGAGCAGAACTCTGCTCCGAGTCC 1751
QY 1681 TTTGACCAATTGGAATTTGCTGACAGCTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1740
Db 1752 TTTGACCAATTGGAATTTGCTGACAGCTTCCGAAACCCCACTGAGAGCTGTGTCT 1811
QY 1741 GTGACAACTGAGGGGCTTTGAGCAGCTCTCTCAGAGTGTGAAGGATGTGGCAG 1800
Db 1812 GTGACAACTGAGGGGCTTTGAGCAGCTCTCTCAGAGGCTTAAAGAGTGTGGCAG 1871
QY 1801 GAATCTTGGGGGATGAGTCTTTCTGACAGACTGCAAGAGTGACTGAGGCTTAC 1860
Db 1872 GAATCTTGGGGGATGAGTCTTTCTGACAGACTGCAAGAGTGAAGAGGCTTAC 1931
QY 1861 AGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
Db 1932 AGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1964

RESULT 8
AX380960
LOCUS AX380960 2929 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 1 from Patent WO0212456.
ACCESSION AX380960
VERSION AX380960.1 GI:19575800
KEYWORDS
SOURCE
ORGANISM
Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
AUTHORS Drucker, D.J., Rosen, C.F. and Iefevre, D.L.
TITLE Ampk-related serine/threonine kinase, designated snark
JOURNAL Patent: WO 0212456-A 1 14-FEB-2002;
1149336 ONTARIO INC. (CA)
FEATURES
source
1..2929
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

Query Match 88.9%; Score 1683.4; DB 6; Length 2929;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

1 ATGAGTGGGTGGCTTAACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 60
 83 ATGAGTGGGTGGCTTAACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 142
 61 TGGGAGAGGCGCCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAACTCTGATGAG 120
 143 AGCGAGAGGCGCCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAACTCTGATGAG 202
 121 AAGCAGCGGTGAGAGGCGGACCATCAAAACAACCTGGGACCGCTACGATTCCTG 180
 203 AAGCAGCGGTGAGAGGCGGACCATCAAAACAACCTGGGACCGCTACGATTCCTG 262
 181 GAGACGCTGGGCAAGGCGACCTACGCGAAGGTGAGAGGCGACGAGAGCTCGGGCGT 240
 263 GAGACGCTGGGCAAGGCGACCTACGCGAAGGTGAGAGGCGACGAGAGCTCGGGACG 322
 241 CTGGTGGCATCAAGTCTCATCAGAGAAACAATAATCAAGATGACAGATCTGTCGAC 300
 323 CTGGTGGCATCAAGTCTCATCAGAGAAACAATAATCAAGATGACAGATCTGTCGAC 382
 301 ATACGAGAGGAGATGAGATCATGTCCTTCACTCAACCCCACTCATTTGCGATCCAT 360
 383 ATACGAGAGGAGATGAGATCATGTCCTTCACTCAACCCCACTCATTTGCGATCCAT 442
 361 GAAGTGTGAGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 420
 443 GAAGTGTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
 421 CTGTATGATTAATCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 480
 503 CTGTATGATTAATCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 562
 481 CGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 563 CGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
 541 AAGCTGAGAAACATCTCTTCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 600
 623 AAGCTGAGAAACATCTCTTCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 682
 601 TCCACCTGTATCAAAAGGCAAGTCTTCAGACGTTCTGTGGAGCCCTCTCTATGCC 660
 683 TCCACCTGTATCAAAAGGCAAGTCTTCAGACGTTCTGTGGAGCCCTCTCTATGCC 742
 661 TCCCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 743 TCCCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 802
 721 GGGGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 803 GGGGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 862
 781 ACACTGTGAGAAACATCTCTTCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 840
 863 ACACTGTGAGAAACATCTCTTCAAGTGAAGCGGCAAGGCTGAGAGCGGACGCAATTTCTTC 922
 841 TGTGGCTGTATCGGCTGGCTGTTAATGATGAGACCCCACTGGGCGCACATGAGAGAT 900
 923 TGTGGCTGTATCGGCTGGCTGTTAATGATGAGACCCCACTGGGCGCACATGAGAGAT 982
 901 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 983 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042
 961 CTGCGTGAAGGAGGAGGAGCCCTAGTGTGATCTTTGGCGGAGCTTCATGAGCGAGTGTGA 1020
 1043 CTGCGTGAAGGAGGAGGAGCCCTAGTGTGATCTTTGGCGGAGCTTCATGAGCGAGTGTGA 1102

1021 CGTGGCTCTCTGGGCGCCCTCTCTGAGAAATGAGCCAGGTGTGAGCTTCTTCAAGCAG 1080
 1103 CGTGGCTCTCTGGGCGCCCTCTCTGAGAAATGAGCCAGGTGTGAGCTTCTTCAAGCAG 1162
 1081 CAGTGTCCGAGAGGTGAGAAACATGATGATGATGATGATGATGATGATGATGATGAT 1140
 1163 CAGTGTCCGAGAGGTGAGAAACATGATGATGATGATGATGATGATGATGATGATGAT 1222
 1141 TCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 1223 TCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
 1201 TCTGGCTGTGAGAGAGAGCTTAACTTCCAGAAAGGCAATTTCTCAAGAAAGGCTCT 1260
 1283 TCTGGCTGTGAGAGAGAGCTTAACTTCCAGAAAGGCAATTTCTCAAGAAAGGCTCT 1342
 1261 ACCTGTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1343 CCTCATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1402
 1321 GGGAGAGCTGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 1403 GGGAGAGCTGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462
 1381 CAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 1463 CAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1522
 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 1523 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1582
 1501 CTCTCTCTCAAGGAGGAGATTTCTCAAACTCAATGAGCAAGTCTTCCGACAGCTTGA 1560
 1583 ---CGCTCTCATGCAAGGAGCATCTCAAACTCAATGAGCAAGTCTTCCGACAGCTTGA 1639
 1561 GAAGGACATACCTCTGACACTTTGAGCTCCCTGAGCAACATGAGCTCTCTCCATCTGCA 1620
 1640 GAAGGACATACCTCTGACACTTTGAGCTCCCTGAGCAACATGAGCTCTCTCCATCTGCA 1699
 1621 GCCCGGCGGAGCGCCCTCTGAGGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 1700 GCCCGGCGGAGCGCCCTCTGAGGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1759
 1681 TTTGACCAATTTGAGCTTGTCTGAGAGCTTTTCCGAAACCCCACTGAGAGGCTGTGTCT 1740
 1760 TTTGACCAATTTGAGCTTGTCTGAGAGCTTTTCCGAAACCCCACTGAGAGGCTGTGTCT 1819
 1741 GTGGAACAACCTGAGGAGGCTGAGAGAGCTTCCCTCAAGAGGTGAGAGAGAGAGAGAG 1800
 1820 GTGGAACAACCTGAGGAGGCTGAGAGAGCTTCCCTCAAGAGGTGAGAGAGAGAGAGAG 1879
 1801 GAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 1880 GAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1939
 1861 AGACAAAGCTTGAAGATTTGCTCAAAAGCTCAGC 1893
 1940 AGACAAAGCTTGAAGATTTGCTCAAAAGCTCAGC 1972

RESULT 9
 AC125887 291762 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
 DEFINITION ***; 13 unordered pieces.
 AC125887
 VERSION AC125887.3 GI:25008671
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 291762)
Muller, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, T., Foster, P., Frazer, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harrey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshauer, L., Louisedge, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawlin, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muthusamy, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwagwu, O., Okunolu, G., Olamposagun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Pu, M., Puozzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smales, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Vals, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D., Holt, R., Smith, H., O., Weinstein, G., and Gibbs, R.A.
Direct Submission
2 (bases 1 to 291762)
Unpublished
Worley, K.C.
Submitted (02-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 291762)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:33269681.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMD
Center clone name: CH230-274J3
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 230853 bases at least Q30
Consensus quality: 233437 bases at least Q20
Estimated insert size: 233445; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 70277: contig of 70277 bp in length
* 70278 70377: gap of unknown length
* 70378 95031: contig of 24654 bp in length
* 95032 95131: gap of unknown length
* 95132 108956: contig of 13825 bp in length
* 108957 109056: gap of unknown length
* 109057 271697: contig of 162641 bp in length
* 271698 271979: gap of unknown length
* 271980 273026: contig of 1229 bp in length
* 273027 273126: gap of unknown length
* 273127 274531: contig of 1405 bp in length
* 274532 274631: gap of unknown length
* 274632 276287: contig of 1656 bp in length
* 276288 276387: gap of unknown length
* 276388 277904: contig of 1517 bp in length
* 277905 278004: gap of unknown length
* 278005 279523: contig of 1519 bp in length
* 279524 279623: gap of unknown length
* 279624 280724: contig of 1101 bp in length
* 280725 280824: gap of unknown length
* 280825 284658: contig of 384 bp in length
* 284659 284758: gap of unknown length
* 284759 287234: contig of 2476 bp in length
* 287235 287334: gap of unknown length
* 287335 291762: contig of 4428 bp in length.
Location/Qualifiers
1..291762
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-274J3"
70378..75596
/note="wgs contig"
95132..96554
/note="wgs contig"
109057..110574
/note="wgs contig"
ORIGIN
Query Match 80.7%; Score 1528.4; DB 2; Length 291762;
Best Local Similarity 99.3%; Pred. No. 0;

Matches 1535; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY	1	ATGAGTGGTGGCTTATCTCCAGGCGCCGAGCGAAGGCTCCCTCGGCTCCGCGCTGGCC	60
Db	274742	ATGAGTGGTGGCTTATCTCCAGGCGCCGAGCGAAGGCTCCCTCGGCTCCGCGCTGGCC	274801
QY	61	TCGAGAGCGCCCGGCGCTGGCGGACGGGCTTCATCAAGTCGCTAAACCTTGATGAAG	120
Db	274802	TCGAGAGAGCGCCCGGCGCTGGCGGACGGGCTTCATCAAGTCGCTAAACCTTGATGAAG	274861
QY	121	AAGCAGCGCGTTAACTCGGCAACATCAACAACACTGCGGACCGGCTTACAGATTCTTG	180
Db	274862	AAGCAGCGCGTTAACTCGGCAACATCAACAACACTGCGGACCGGCTTACAGATTCTTG	274921
QY	181	GAGACGCTGGGCAAGGCGACCTACGGGAAGGTAAAGGACGAGAGAGCTCGGAGCGT	240
Db	274922	GAGACCTGGGCAAGGCGACCTACGGGAAGGTAAAGGACGAGAGAGCTCGGAGCGT	274981
QY	241	CTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGACGAGATCTTGTCAC	300
Db	274982	CTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGACGAGATCTTGTCAC	275041
QY	301	ATACGGAGGGAGATTGAGATCAATGTCTTCACTCAACACCCCACTCATATGCGATCCAT	360
Db	275042	ATACGGAGGGAGATTGAGATCAATGTCTTCACTCAACACCCCACTCATATGCGATCCAT	275101
QY	361	GAAGTGTTTGAGAAATAGCAGCAAGATTGTATTTGTCAATGAGATGATCCAGCCGAGCGAT	420
Db	275102	GAAGTGTTTGAGAAATAGCAGCAAGATTGTATTTGTCAATGAGATGATCCAGCCGAGCGAT	275161
QY	421	CTGTATGATTATCATGATGATGAGCGGCGCACGGCTGATGAGCGGGAACCCACAGCATTTCTTC	480
Db	275162	CTGTATGATTATCATGATGATGAGCGGCGCACGGCTGATGAGCGGGAACCCACAGCATTTCTTC	275221
QY	481	CGACGAGATCGATCTGCGCCCTGACATCTGCGCACACAGGGAGTGGTTACACGAGATCTC	540
Db	275222	CGACGAGATCGATCTGCGCCCTGACATCTGCGCACACAGGGAGTGGTTACACAGAGATCTC	275281
QY	541	AAGCTGAGAAAATCATCTTCTAGATGCGCAATGGAACATCAAGATTGCTGACTTTGGCCTC	600
Db	275282	AAGCTGAGAAAATCATCTTCTAGATGCGCAATGGAACATCAAGATTGCTGACTTTGGCCTC	275341
QY	601	TCCAACTCTGTACCACAAAGGCAAGTTCCTCCAGAGCTTCTGTGGAGCCCTCTCTACGCC	660
Db	275342	TCCAACTCTGTACCACAAAGGCAAGTTCCTCCAGAGCTTCTGTGGAGCCCTCTCTACGCC	275401
QY	661	TCGCGCTGAGATGTACACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGTCTCTG	720
Db	275402	TCGCGCTGAGATGTACACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGTCTCTG	275461
QY	721	GGCGTCTCTGTATACATCTGTGTGATGAGCACCATGACCCTTTTGAACGGGACGATCTATAA	780
Db	275462	GGCGTCTCTGTATACATCTGTGTGATGAGCACCATGACCCTTTTGAACGGGACGATCTATAA	275521
QY	781	ACACGTGTGAAGCAATATCATGATACGGGGCTTACCGTGAGCCCGCCAAAGCCCTCCATATGCC	840
Db	275522	ACACGTGTGAAGCAATATCATGATACGGGGCTTACCGTGAGCCCGCCAAAGCCCTCCATATGCC	275581
QY	841	TGTGGCTCATGATCCGGTGGCTGTTAATGTGTGAACCCACCCGCTCGGGCCACACTGTAGAGAT	900
Db	275582	TGTGGCTCATGATCCGGTGGCTGTTAATGTGTGAACCCACCCGCTCGGGCCACACTGTAGAGAT	275641
QY	901	GTAACCAATCATTTGTGGGTCAACTGGGGTTACACCAACGAGATCGGGGAAACGAAAGCC	960
Db	275642	GTAACCAATCATTTGTGGGTCAACTGGGGTTACACCAACGAGATCGGGGAAACGAAAGCC	275701
QY	961	CTGCGTGAAGGTGGGCAACCTATGTGGTACTTTTGGCCGAGCCCTCAATGGCGGACTGGTTA	1020
Db	275702	CTGCGTGAAGGTGGGCAACCTATGTGGTACTTTTGGCCGAGCCCTCAATGGCGGACTGGTTA	275761
QY	1021	CGTGGCTCTCGCGGCCCTCTCTCGAGAAATGAGCCAGAGGTGTGACGCTTTTCAAGCAG	1080
Db	275762	CGTGGCTCTCGCGGCCCTCTCTCGAGAAATGAGCCAGAGGTGTGACGCTTTTCAAGCAG	275821

Qy	1081	CACGTCCGGAGGGTGAAGCACTTACTCTGGGCTGGAGCGCAACTTCTTTAAGAA	1140
Dp	275822	CACGTCCGGAGGGTGAAGCACTTACTCTGGGCTGGAGCGCAACTTCTTTAAGAA	275881
Qy	1141	TCCGGAAGAGAAATGACATGCTCAAAATCTGCAGAGTGACCCGGCTGAGATACCTT	1200
Dp	275882	TCCGGAAGAGAGAAATGACATGCTCAAAATCTGCAGAGTGACCCGGCTGAGATACCTT	275941
Qy	1201	TCTCGCCCTGGCAAGAGCAAGCCTTAAGCTTCCGAAGGCAATCTCAAGAAAAAGTCTCT	1260
Dp	275942	TCTCGCTCTGGCAAGAGCAAGCCTTAAGCTTCCGAAGGCAATCTCAAGAAAAAGTCTCT	276001
Qy	1261	ACCTGTCAGGGAGGTACAGAGAGACCCTCAGAACTCAGACCGGTGCTGATACCTCA	1320
Dp	276002	ACCTGTCAGGGAGGTACAGAGAGACCCTCAGAACTCAGACCGGTGCTGATACCTCA	276061
Qy	1321	GGGCAAGCTTGTCCCTGCTGTATCCCTGCTCCCAAGAAAAAGCATCTTTAAGAAAGTCTCA	1380
Dp	276062	GGGCAAGCTTGTCCCTGCTGTATCCCTGCTCCCAAGAAAAAGCATCTTTAAGAAAGTCTCA	276121
Qy	1381	CAGCGTGAATCTGGTTACTACTCTCTCAGAGCCAGCGAATCTGGGGAACTCTTAGAC	1440
Dp	276122	CAGCGTGAATCTGGTTACTACTCTCTCAGAGCCAGCGAATCTGGGGAACTCTTAGAC	276181
Qy	1441	GCCAGTGAATGTTTGTGTGAGTGGGAGCCCGTGGAGCAGAAATCTTCACAGGCTTCAGGG	1500
Dp	276182	GCCAGTGAATGTTTGTGTGAGTGGGAGCCCGTGGAGCAGAAATCTTCACAGGCTTCAGGG	276241
Qy	1501	CTCTCTCCCAACCGCAAGGCAATTCGAAATCAATGGCAAGTTTCT	1546
Dp	276242	CTCTCTCCCAACCGCAAGGCAATTCGAAATCAATGGCAAGTTTCT	276287

RESULT	10
AX407118	
LOCUS	AX407118
DEFINITION	AX407118 1884 bp DNA linear PAT 14-JUN-2002
ACCESSION	Sequence 3 from Patent WO0224921.
VERSION	AX407118
KEYWORDS	AX407118.1 GI:21439883
SOURCE	.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1
AUTHORS	Curtis,R.A. and Galvin,K.M.
TITLE	3700, a novel human protein kinase and uses therefor
JOURNAL	Patent: WO 0224921-A 3 28-MAR-2002;
FEATURES	Millennium Pharmaceuticals, Inc. (US) Location/Qualifiers 1..1884 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
source	
ORIGIN	

	Query Match	Best Local Similarity	Score	DB	Length
	Matches 1565; Conservative	71.9%;	1360.2;	6;	1884;
		Pred. No. 0;			
	Mismatches 273; Indels 21; Gaps 2;				
QY	47	CCTCCGCCCTTGCCCTCGAGAGCGCCCGCGCTGCGGACGGGCTATCAAGTCGCTA	106		
Db	35	CCACTCCCTCGGCGCGCAGAGCTAGCCCGCGCGCTGGGGAAGGCTATCAAGTCGCCCA	94		
QY	107	AACTCTGATGAGAGACAGCGCGGTGAAGCGGACCATCACAAAACAACCTGCGGCAAC	166		
Db	95	AGCCCTTAATGAAGAGACAGCGGTGAAGCGGACCAACAACAACCTGCGGCAAC	154		
QY	167	GCTACGAGTTCCTGGAAGCGCTGGGCAAGGGCACTTACGGGAAGGTAAAGAGGCGACAG	226		
Db	155	GCTACGAGTTCCTGGAAGCGCTGGGCAAGGGCACTTACGGGAAGGTAAAGAGGCGG	214		

227 AGAGCTCGGGGCGTGTGGCCATCAAGTCCATCAGAAAGACAAATATCAAGATGAC 286
215 AAGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCAGAAAGACAAATATCAAGATGAC 274
287 AGAGTGTGTGCAATACGAGGAGGAGATTGAGATCAATGTCTTCACTCAACACCCGACA 346
275 AAGATCTGATGCAATACGAGGAGGAGATTGAGATCAATGTCTTCACTCAACACCCGACA 334
347 TCATTGCCATCCATGAAAGTGTGAGAAATGACAGCAAGATTGATGTGATGAGATG 406
335 TCATTGCCATCCATGAAAGTGTGAGAAACAGCAGCAAGATGTGATGTGATGAGATG 394
407 CCAGCGGAGCGCATCTGTATGATTATCACTAGTAGCGGCGACCGGCTGAGTAGCGGAG 466
395 CCAAGCGGCGGCGACCTTTATGACTTAATCAAGCAAGCGGCGACAGCTCAGTAGCGGAG 454
467 CCAGGCAATTTCTTCCAGAGATGTGTGCTGCTGCACTACTGCAAGAGGAGATG 526
455 CTAGGCAATTTCTTCCGAGAGATGTGTGCTGCTGCACTACTGCAAGAGATG 514
527 TTCACCGAATCTCAAGCTGAGAAACATCTTCAATGCGAATGGAACATCAAGATG 586
515 TCACCGAATCTCAAGCTGAGAAACATCTTCAATGCGAATGGAACATCAAGATG 574
587 CTGACTTTGAGCTTCCAACTGTACCAAGAGGAGTTCCCTCAGAGCTGTGTGAG 646
575 CTGACTTTGAGCTTCCAACTGTACCAAGAGGAGTTCCCTCAGAGCTGTGTGAG 634
647 GCCCTCTCAAGCTGCTGAGATGATCAACGAGGAGCTTATGTGTGAGGCGCAGAGT 706
635 GCCCTCTCAAGCTGCTGAGATGATGATCAACGAGGAGCTTATGTGTGAGGCGCAGAGT 694
707 ACAGCTGTGTCTGTGGGCTTCTCTGTACATCTGTGTGATGACACCATGCTTTGAG 766
695 ACAGCTGTGTCTGTGGGCTTCTCTGTACATCTGTGTGATGACACCATGCTTTGAG 754
767 GCGAGGATCATAAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCGGCGCA 826
755 GCGAGGATCATAAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCGGCGCA 814
827 AGCCGCTCGATGCTGTGGCTGATCCGCTGCTGTTAATGTGAAACCCGACCGCTGAG 886
815 AACCCCTGATGCTGTGGCTGATCCGCTGCTGTTAATGTGAAACCCGACCGCTGAG 874
887 CCAACTGAGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 946
875 CCAACTGAGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 934
947 GCGAGGATCATAAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCGGCGCA 1006
935 GCGAGGATCATAAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCGGCGCA 994
1007 TGGCGGACTGTTCAGTGTGCTGCTGCGGCGGCTTCTGAGAGATGAGGCAAGTGTGA 1066
995 TGGCGGACTGTTCAGTGTGCTGCTGCGGCGGCTTCTGAGAGATGAGGCAAGTGTGA 1054
1067 GCTTCTTCAAGGAGCATGACCTGTGGGAGGAGCACTGTAAGTGTGAGGCTGAGGCAAG 1126
1055 GCTTCTTCAAGGAGCATGACCTGTGGGAGGAGCACTGTAAGTGTGAGGCTGAGGCAAG 1114
1127 ATTCTCTTAAAGATCCGAGAGGAGATGACATGCTGTAAGTGTGAGGCTGAGGCAAG 1186
1115 ATTCTCTTAAAGATCCGAGAGGAGATGACATGCTGTAAGTGTGAGGCTGAGGCAAG 1174
1187 CTGAGGATCACTTCTGCGGCTGAGAGGAGGCTTAAAGTGTGAGGCAAGGCTGTA 1246
1175 CTGAGGATCACTTCTGCGGCTGAGAGGAGGCTTAAAGTGTGAGGCAAGGCTGTA 1234
1247 AAAAAAGTCTTCACTGTGTAAGGAGGATGAGAGGAGGCTTCAAGAGCTTCAAGCG 1306
1235 AAAAAAGTCTTCACTGTGTAAGGAGGATGAGAGGAGGCTTCAAGAGCTTCAAGCG 1294
1307 TGCCTGATCTCAAGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366

1295 TCCCTCGAGCCGAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345
1367 TTAAGAGTCTGACACAGCGTGAATCTGTACTACTCTCTCAGAGCCGAGGAGTCT 1426
1346 TCAGAGAGCCCGAGAGCGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1405
1427 GGGAACTCTTAAAGCGCAAGT 1486
1406 GGGAGCTCTTAAAGCGCAAGT 1465
1487 CACAGGCTTCAAGGAGCTCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1546
1466 CCGAGCTTCAAGGAGCTCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1525
1547 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1606
1526 CCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1585
1607 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
1586 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1645
1667 TGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1726
1646 TGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1705
1727 GGGGCTGT 1780
1706 GGGGCTGT 1765
1781 -----GTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1834
1766 GAGGCTGT 1825
1835 ACTGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1893
1826 ACTGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1884

RESULT 11
AX407116
LOCUS AX407116 3353 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0224921.
ACCESSION AX407116
VERSION AX407116.1 GI:21439882
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1
Curtis, R.A. and Galvin, K.M.
3700, a novel human protein kinase and uses therefor
Patent: WO 0224921-A 1 28-MAR-2002;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..3353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 71.9%; Score 1360.2; DB 6; Length 3353;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

47 CCTCGGCTGTGCTGAGAGCGCCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 106
191 CCATTCCTTCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 250
107 AACCTGTATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 166

Db 251 AGCCCTTAATGAAGAAGAGGCGGTGAAGCGGACCAACCAAGCAAACTGCGGACCC 310
Qy 167 GCTACAGATTCCTGAGAGCGCTGGGCAAGGSCACTACGGGAAGTGAAGAAGCAGAG 226
Db 311 GCTACAGATTCCTGAGAGCGCTGGGCAAGGSCACTACGGGAAGTGAAGAAGCAGAG 370
Qy 227 AAGAGCTGGGGCGCTGCTGGGCGCATCAAGTCACAGAAAGCAAAATCAAAAGATGAG 286
Db 371 AAGAGCTGGGGCGCTGCTGGGCGCATCAAGTCACAGAAAGCAAAATCAAAAGATGAG 430
Qy 287 AAGATCTGCTGCAATACGAGAGGAGATTTGAGATCATGCTTTTACCTCAACACCCCA 346
Db 431 AAGATCTGATGCAATACGAGAGGAGATTTGAGATCATGCTTTTACCTCAACACCCCA 490
Qy 347 TCATTCGCAATCCATGAAGTGTGAAGATTAAGCAGAAAGTGTGATTTGATGAGATG 406
Db 491 TCATTCGCAATCCATGAAGTGTGAAGATTAAGCAGAAAGTGTGATTTGATGAGATG 550
Qy 407 CCAGCCGAGGCGCATCTGATGATTAATCAATGAGCGGCGCACCGCTGATGAGCGGAG 466
Db 551 CCAGCCGAGGCGCATCTTATGATCAATGAGCGGCGGAGGAGCTCAAGTGAAGCGGAG 610
Qy 467 CCAAGCATTTCTTCCGACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
Db 611 CTAAGCATTTCTTCCGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
Qy 527 TTACCGGATCTCAAGCTGGAAGAAATCTTATGATGCAATGGAATCAATGAATG 586
Db 671 TCACCGGATCTCAAGCTGGAAGAAATCTTATGATGCAATGGAATCAATGAATG 730
Qy 587 CTGACTTTGGCCCTCTCAACCTGACCAACAAAGCAAGTCTCTCAGACCTTCTGAG 646
Db 731 CTGACTTTGGCCCTCTCAACCTGACCAACAAAGCAAGTCTCTCAGACCTTCTGAG 790
Qy 647 GCCCTCTCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
Db 791 GCCCTCTCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
Qy 707 ACAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 766
Db 851 ACAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
Qy 767 GGCAGATCATAAACAAGTGTGAAGCAATCAATCAAGGAGCTTACCTGAGCCGCA 826
Db 911 GGCAGATCATAAACAAGTGTGAAGCAATCAATCAAGGAGCTTACCTGAGCCGCA 970
Qy 827 AGCCGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 886
Db 971 AACCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
Qy 887 CCAACTGAGAGATGTAAGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946
Db 1031 CCAACTGAGAGATGTAAGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
Qy 947 GGGAAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 1091 GAGAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150
Qy 1007 TGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
Db 1151 TGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Qy 1067 GCTTCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
Db 1211 GCTTCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
Qy 1127 ATTCTCTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186
Db 1271 ATTCTCTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
Qy 1187 CTGAGATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
Db 1331 CTGATGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390

Qy 1247 AGAAAAAGTCTCTACTCTGCTGAGGGAGATACAGAGGAGCCCTCAGAACTCAGACCGG 1306
Db 1391 AGAAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
Qy 1307 TGGCTGATCTCAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366
Db 1451 TGGCTGATCTCAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1501
Qy 1367 TTAAGAGTCTGAGCAGCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426
Db 1502 TCAAGAGGCTGAGCAGCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
Qy 1427 GGGAACTCTTAAAGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1486
Db 1562 GGGAGCTTTGAGCGCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621
Qy 1487 CAGAGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1546
Db 1622 CGCAAGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1681
Qy 1547 CCGCAGCGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606
Db 1682 CCGCAGCGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741
Qy 1607 CCGCAGCGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1666
Db 1742 CCGCAGCGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1801
Qy 1667 TGTCTCCAGATCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726
Db 1802 TGTCTCCAGATCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
Qy 1727 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1780
Db 1862 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
Qy 1781 -----GTCTGAAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1834
Db 1922 GAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1981
Qy 1835 ACTGCAGAGGAGTGAAGGAGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1893
Db 1982 ACTGCAGAGGAGTGAAGGAGCTTAAAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2040

RESULT 12
AX642961 3360 bp DNA linear PAT 24-FBB-2003
LOCUS AX642961
DEFINITION Sequence 38 from Patent WO01096547.
ACCESSION AX642961
VERSION AX642961.1 GI:28550104
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Yue, H., Lal, P., Bandman, O., Borowsky, M., Au-Young, J., Lu, Y.,
Gandhi, A.R., Tribouley, C.M., Walla, N., Yao, M.G., Lu, D.A.,
Greenwald, S.R., Ramkumar, J., Griffin, J.A., Kearney, L., Buford, N.,
Nguyen, D.B., Tang, Y.T., Baughn, M.R., He, A., Thornton, M.,
Hafalia, A., Patterson, C., Gururajan, R., Lo, T.P., Khan, F.,
Recipon, S.A., Azimzal, Y., Policky, J.L., Ding, L., Grether, M.,
Elliot, V.S., Thangavelu, K., Batra, S. and Ison, C.H.
HUMAIN KINASES
JOURNAL
Patent: WO 01096547-A 38 20-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1.3360
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN /note="Incyte ID No: 4841542GB1"

Query Match 71.7%; Score 1357.6; DB 6; Length 3360;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;

```
QY 32 GCCAGGCTCCCTGGGCTCCGCTCCGAGAGCGCCCGCTGGCGGACGAGC 91
DB GCGCGGCTCCGCGCCCACTCCCTCGCGCAGAGCTAGCCCGCCCTGGCGGAGGGC 157
QY 92 TCATCAAGTGCCTAAACTCTGATGAAAGAGCGGCGTGAAGCGGACCATCAAAAC 151
DB 158 TGATCAAGTGCCTAAAGCTCTGATGAAAGAGCGGCGTGAAGCGGACCATCAAGC 217
QY 152 ACAACCTGCGGCAACCGCTACGAGTCTCTGAGACGCTGGGCAAGGCACTTACGGAGG 211
DB 218 ACAACCTGCGGCAACCGCTACGAGTCTCTGAGACGCTGGGCAAGGCACTTACGGAGG 277
QY 212 TGAAGAAAGGACGAGAGAGTCCGGGCGCTGCTGGGCAATCAAGTCAAGAGAGACA 271
DB 278 TGAAGAAAGGCGGAGAGAGTCCGGGCGCTGCTGGGCAATCAAGTCAAGAGAGACA 337
QY 272 AATCAAGATGAGAGAGATCTGCTGCAATACGAGAGAGATTTAGATCATGCTTTAC 331
DB 338 AATCAAGATGAGAGAGATCTGCTGCAATACGAGAGAGATTTAGATCATGCTTTAC 397
QY 332 TCAACCAACCCCAATCATTTCCATTCATGAGTGTGTAAGATAGCAGCAAGATTGTA 391
DB 398 TCAACCAACCCCAATCATTTCCATTCATGAGTGTGTAAGATAGCAGCAAGATTGTA 457
QY 392 TTGTCAAGAGTATGACAGCGGAGCGATCTGATGATTTACATCAAGTGAAGCGCACGCG 451
DB 458 TGTCAAGAGTATGACAGCGGAGCGATCTGATGATTTACATCAAGTGAAGCGCACGCG 517
QY 452 TGAAGTGAAGGAGAGCGCAGGCAATTTCTTCCGACAGATGCTGTCGCGCTGCACTAGTCC 511
DB 518 TGAAGTGAAGGAGAGCGCAGGCAATTTCTTCCGACAGATGCTGTCGCGCTGCACTAGTCC 577
QY 512 ACCAGAACGGAATGCTTACCGAGATCTCAAGCTGGAACCATCTTCTAGATGCCAATG 571
DB 578 ATCAGAACGGAATGCTTACCGAGATCTCAAGCTGGAACCATCTTCTAGATGCCAATG 637
QY 572 GAAACATCAAGATGCTGATTTGGGCTCTCAACCTGTAACCAAGGCAATGCTTCTCC 631
DB 638 GGAATATCAAGATGCTGATTTGGGCTCTCAACCTGTAACCAAGGCAATGCTTCTCC 697
QY 632 AGACGTTCTGTGGAGCGCTCTCTACGCGCTGCTGAGATGCAACGCGGAGCCCTATG 691
DB 698 AGACGTTCTGTGGAGCGCTCTCTACGCGCTGCTGAGATGCAACGCGGAGCCCTATG 757
QY 692 TGGGCGCAGAGTGAACAGTGTCTCTGGGCGTTCTCTGTAACATCTGTGATGCAATG 751
DB 758 CAGGCGCAGAGTGAACAGTGTCTCTGGGCGTTCTCTGTAACATCTGTGATGCAATG 817
QY 752 CCATGCGCTTTTGAACGCGGAGAGATCAATAACATCTGTGAAGCAATCAAGTCAAGG 811
DB 818 CCATGCGCTTTTGAATGGGCAATGAACATTAAGTCTTAAGTGAAGCAATCAAGTCAAGG 877
QY 812 ACCGTGAGCGCGCCAGAGCGCTGCTGAGTGGCGCTGATCCGCTGAGCTGTTAATGTTGA 871
DB 878 ACCGTGAGCGCGCCAGAGCGCTGCTGAGTGGCGCTGATCCGCTGAGCTGTTAATGTTGA 937
QY 872 ACCCCACCGCTGGGCGCACATGAGAGATGTAACGATCAATTTGTGGTCAACTGGGGTT 931
DB 938 ACCCCACCGCGCGGCGCACCTGAGAGATGTTGGCGATCATCTGTGGTCAACTGGGGCT 997
QY 932 AACACCAACCGGAGTCCGGGAGACAGAAAGCCCTGCTGTAAGGGTGGGCACTTATGTTGACT 991
DB 998 AACGCAACCGGAGTCCGGGAGACAGAAAGCTCCGATATAGGGTGGGCACTTATGTTGACT 1057
QY 992 TTGGCGGCGCTTCATGCGGAGATGTTTGGTCTCTCGGCGCCCTCTGGAGAAAG 1051
```

```
DB 1058 CTGCCCCGCGCTTCATGCGTGAAGTCTCCGCGCTTCTCCGCCCCCTCTCTGAGAAATG 1117
QY 1052 GAGCCAAAGTGTGACAGCTTCTTCAAGACAGACAGTGCAGGAGGTGAAGCACTGATCTG 1111
DB 1118 GGGCCAAAGTGTGACAGCTTCTTCAAGACAGACAGTGCAGGAGGTGAAGCACTGATCTG 1177
QY 1112 GGTGAGAGGGGCAACATTTCTTAAAGATCCCGAAAGAAATGACATGCTCAAAATC 1171
DB 1178 GCTTGAAGGCGCCAGCATTTGCTCAAGAAATCCCGCAAGAGATGACATGCGCCAGTCTTC 1237
QY 1172 TGCAGGTGACCCGCGTGAAGATCTCTTCTGCGCTGCGCAAGACAGCTTAACTTC 1231
DB 1238 TCCACGTGACAGGCTGATGACATGCGCAATGCGCTGCGCAAGACAGCTTAACTTC 1297
QY 1232 CGAAAGCATTTCTCAAGAAAGATCTCTTAACTCTGTCAGGAGAGTACAGAGACCTTC 1291
DB 1298 CAAGGCGCATTTCTCAAGAAAGATGTCAGCTCTGCAAGAGGAGTACAGAGACCTTC 1357
QY 1292 AGGAATCTGACCGGCTGCTGATATCTCAAGGAGAGCTGCTCCCTGCTATCCCTGCTCC 1351
DB 1358 CGAGCTCAGCCCAATCCCTGCGAGCCAGGCGAGGCTGCC-----CCCTGCTCC 1408
QY 1352 CAAGAAAGCATCTTAAAGATCTGCAAGGCTGCAAGGCTGCAAGCTTAACTCTCTCCAG 1411
DB 1409 CCAAGAGGCGATTTCTCAAGAAAGCTCCGACAGGCGAGATCTGCTATCTCTCTCCG 1468
QY 1412 AGCCAGAGAGTCTGGGGAATCTTAAAGCAAGTATGTTGATGAGGAGACCCG 1471
DB 1469 AGCCAGTGAATCTGGGAGATCTTGAAGCAAGGAGAGCTGTTGATGAGGAGATCCA 1528
QY 1472 TGAAGCAAGTCTCCAGAGGCTTCAAGGCTCTCTCTCAACGCAAGGAGATTTCTCAAC 1531
DB 1529 AGAGAGCAAGGCTCCGCAAGGCTTCAAGGCTCTCTCTCAAGGAGATTTCTCAAC 1588
QY 1532 TCAATGCAAGTCTCCGCGACAGCTTGAAGAGCACTACCCCTGACATTTGGCTCC 1591
DB 1589 TCAATGCAAGTCTCCGCGACAGCTTGAAGTGCAGGCTCCGACACATTTGGCTCC 1648
QY 1592 TGAACCACTGCGCTCTCCATCTCTGACAGCCCGCCAGCGCCCTGAGGAGCTGTA 1651
DB 1649 TGAATGAATCTGCGCCACCTGCGCCCTGCGCCCGGCGCAGCGGCTGTA 1708
QY 1652 GTGAGACAGCATCTGCTCTCTGAGTCTTGAACCAATGGAATTTGCTGAAAGTCTTC 1711
DB 1709 GCGAGACAGCATCTGCTCTCTGAGTCTTGAACCAATGGAATTTGCTGAAAGTCTTC 1768
QY 1712 CGAAACCCACAGAGGCGCTGCTGTAAGCAACTGAGAGGCGCTTGAAGAGCTTC 1771
DB 1769 CAGAGCCCACTGCGGAGCTGCTGTAAGCAACTGAGAGGCGCTTGAAGAGCTTC 1828
QY 1772 CCTCAGAAAG-----GTCTGAAGCAGATGCTGAGAGAAATCTTGGGAGATAGCT 1819
DB 1829 CCTCAGAGGCGCTTGAAGAGCTCTGAGGCGCTGCGGAGAGATCTTGGGAGAGCT 1888
QY 1820 GCTTTTCTTGAACAGCTGCAAGAGTGAATGCAAGCTTACAGACAGCTTGAAGATCT 1879
DB 1889 GCTTTTCTTGAACAGCTGCAAGAGTGAATGCAAGCTTACAGACAGCTTGAAGATCT 1948
QY 1880 GCTCAAGCTCAGC 1893
DB 1949 GCTCAAGCTCAGC 1962
```

RESULT 13
CQ714327 1887 bp DNA linear PAT 03-FEB-2004
DEFINITION
Sequence 261 from Patent WO20068579.
ACCESSION
CQ714327
VERSION
CQ714327.1 GI:42275184
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

LOCUS C0782778 3395 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2918 from Patent EP1396543.
ACCESSION C0782778
VERSION C0782778.1 GI:45502721
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Okada, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2918 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source
1. .3395
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
118. .2004
/note="unnamed protein product"
/codon_start=1
/protein_id="CAFA6008.1"
/db_xref="GI:45502722"
/translation="MESLVFARRSGTPPSAELARPLAELIKSPKPLMKQAVKHH
HKNLHRYEFLETKGYKVKARBSGRVLAISKIKDIKDEDMHMRLE
INSLNHPHIAIHEVENSISKIVIMEYSRDLVISERGLSEBRLHPPRIIV
SAVYCHQNRVHRDLKLENIILADNGNIKIADPGSNLHOGKFLIOTFGSLVAP
EIVNGKPYGPEVDSWSIGVLIYILVHGTPMPDGHDKILVNOISGAYREPSPDA
CGILRPLVWNPTRRATLEVDVASHMWNKGYATRVBQEAPEHGHGSDASVAD
WLRSSRPLENGAKVCSFFKQHPGGSTTPGLERQHSIKSRKNDMAQSHSTJA
DIDAHKPKSNLKPGLIKKVSASAEVQEDPPELSPIDAPPGQALPLPKGILK
KPRORSGYSSPESPESELIDAGVFPVSGDEKQKPPQASGLILHKGILILNKP
SCFALRLAAPTTFGSLDELAPPLPARASPSGAVSEDTISSESDDLPERLEBP
PLRGVSVNLTGLEBPSBSPGSCLRMRKQDPLDGSCEBTLDCQVATRYQALNVC
SKLT"

ORIGIN
Query Match 71.7%; Score 1357; DB 6; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCTCCGCGCTGGCTGGAGAGCGCCGCGCTGGCGAGCGGGCTTCATCAAGTCGCTTA 106
DB 152 CCACTCCCTCGCGCGGAGAGCTAGCCGCGCGCTGGCGAGAGGGCTGATCAAGTCGCCCA 211
QY 107 AACCTCGATGAAG 166
DB 212 AGCCCTAAATGAAG 271
QY 167 GCTACGAGTTCCTGGAGAGCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 226
DB 272 GCTACGAGTTCCTGGAGAGCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
QY 227 AGAGCTCGGGCGCTGGTGGCCATCAAGTCCATCAGAGAGAGAGAGAGAGAGAGAGAGAG 286
DB 332 AGAGCTCGGGCGCTGGTGGCCATCAAGTCCATCAGAGAGAGAGAGAGAGAGAGAGAGAG 391
QY 287 AGGATTCGTGCAATACGAG 346
DB 392 AAGATTCGTGCAATACGAG 451
QY 347 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
DB 452 TCATTGCCATCCATGAAGTGTGAGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511
QY 407 CCAAGCGAG 466
DB 512 CCAAGCGAG 571
QY 467 CCAAGGATTTCTTCGACAGATGTGTGCTGCTGCACTATGCTCCACAGAGAGAGAGATG 526

DB 572 CTAGGCAATTTCTTCGGCAGATGTCTGCGCTGCACTATGCTCAACAGAGAGAGAGAG 631
QY 527 TTCACGAGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586
DB 632 TCACCGAGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 691
QY 587 CTGACTTTGGCTCTTCAACCTGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646
DB 692 CTGACTTTGGCTCTTCAACCTGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
QY 647 GCCCTCTACGCTGCGCTGAGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
DB 752 GCCCTCTACGCTGCGCTGAGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
QY 707 ACAGCTGTCTCTGGCGCTTCTCTGTAATCATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 766
DB 812 ACAGCTGTCTCTGGCGCTTCTCTGTAATCATCTGCTGAGAGAGAGAGAGAGAGAGAGAG 871
QY 767 GCGAGAGATCAATAACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
DB 872 GCGATGACCATTAAGATCTATGAAACAGATCAGCAACGAGGCTTACCGAGAGAGAGAGAG 931
QY 827 AGCCGCTGATGCTGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
DB 932 AACCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
QY 887 CCACATGAGAGAGATGAG 946
DB 992 CCACTCTGAGAGATGAG 1051
QY 947 GCGAG 1006
DB 1052 GAG 1111
QY 1007 TGGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
DB 1112 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171
QY 1067 GCTTCTTCAAG 1126
DB 1172 GCTTCTTCAAG 1231
QY 1127 ATTCTTAAAGATGCTGAG 1186
DB 1232 ATTCTTAAAGATGCTGAG 1291
QY 1187 CTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
DB 1292 CTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1351
QY 1247 AG 1306
DB 1352 AG 1411
QY 1307 TGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366
DB 1412 TGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
QY 1367 TTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426
DB 1463 TTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522
QY 1427 GGAAGCTTGAAG 1486
DB 1523 GGAAGCTTGAAG 1582
QY 1487 CACAGGCTTCAAGGCTCTCTCTCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1546
DB 1583 CCAAGGCTTCAAGGCTCTCTCTCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1642
QY 1547 CCGGCAAGGCTTGAAG 1606

Db 1292 CTGATGACACTGCGCATCGCCCTGGCAAGACCACTCAAGCTGCGCAAGGCGATTCTCA 1351
 QY 1247 AGAAAAGTCTCTACTCTGTCAGGGAGGTATACAGAGGACCTTCAGGAACTCAGACCGG 1306
 Db 1352 AGAAGAGGTGTGAGCTTGCAGAAAGGGGTACAGAGGACCTCCGAGCTCAGGCCAA 1411
 QY 1307 TGCCTGATACTCCAGGGCAGCCTGTCTGTATCCCTGCTCCCAAGGAAGGCATCC 1366
 Db 1412 TCCCTCGAGCCCGAGGCGAGGCTGCC-----CCTGCTCCCAAGAGGGCATTC 1462
 QY 1367 TTAAGAAGTCTCGACAGCCGTGAATCTGTTACTACTCTCCAGAGCCCGAGGCTTG 1426
 Db 1463 TCAGAAGGCCCGCAGCGCGAGTCTGCTACTACTCTCTCCGAGCCCGAGTAACTTG 1522
 QY 1427 GGGAACTCTTAAAGCCGAGTATGTGTTTGTGAGTGGGAGCCCGTGGAGCAGAAGTTC 1486
 Db 1523 GGGAGCTCTTGGACGCAAGGCGAGCTGTTGTGAGTGGGATCCCAAGAGCAGAAGCTTC 1582
 QY 1487 CACAGGCTTCAGGGCTCTCTCTCACCGCAAGGGCATTCCAAACTCAATGGCAAGTTC 1546
 Db 1583 CGAAGCTTCAGGGCTGCTCTCTCTCAATCGCAAGGCGATCTCAAACTCAATGGCAAGTTC 1642
 QY 1547 CCGGCAAGCTTAAAGGCACTACCCCTAGCACTTTGCTCCTGACCAACTGAGCTT 1606
 Db 1643 CCGAGACAGCTTGAAGCTGCGGCGCCCACTTCGAGCTCCTGAGTAACTGCGCC 1702
 QY 1607 CCTCCCATCTGACAGCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGACAGCATCC 1666
 Db 1703 CACCTGCGCCCTGGCGCCCGGCCAGCGCATCTCAGGGGCTGTGAGTGAAGACAGCATCC 1762
 QY 1667 TGTCTCCGAGTCTTGAACCAATTGGAATTTGCTGAAGCTTCCCGAAACCCCACTGA 1726
 Db 1763 TGTCTCTGAAGTCTTGAACCACTGGAATTTGCTGAACGAGCTCCAGAGCCCACTGC 1822
 QY 1727 GGGGCTGTGTGTGTGACCAACCTGAGGGGCTTGAAGCAGCTCCCTCAGAAG----- 1780
 Db 1823 GGGGCTGTGTGTGTGACCAACCTGAGGGGCTTGAAGAGGCCCTCAGAGGGCCCTG 1882
 QY 1781 -----GTCTGAAGCGATGTGTGAGAGATCTTTGGGGGATAGCTGCTTTCTTGAAG 1834
 Db 1883 GAAGCTGCTGAGGCGCTGCGCGGAGATCTTTGGGGGACAGCTGCTTTCCCTGACAG 1942
 QY 1835 ACTGCCAAGAGGTGACTGACGCTTACAGACAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893
 Db 1943 ACTGCCAAGAGGTGACAGGAGCTTACCGACAGGCACTGAGGATCTGCTCAAAAGCTCAC 2001

Search completed: May 15, 2005, 22:46:06
 Job time : 8121.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 16:05:16 ; Search time 9307.58 Seconds

(without alignments)
11868.013 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2302
Sequence: 1 cactagtcgacccaagaat.....aaaaaaaaaaaaaaaaa 2302

Scoring table: IDENTITY NUC
Gapop 10%0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_est8.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2865.8	98.8	2899	3 AK004737	AK004737 Mus muscu
2	2865.8	98.8	3098	3 AK033672	AK033672 Mus muscu
3	2804.4	96.6	2869	3 AK034082	AK034082 Mus muscu
4	1776	61.2	1896	9 AY399041	AY399041 Mus muscu
5	1374.2	47.4	3443	3 HSM801859	AL136891 Homo sapi
6	1264	43.6	1887	9 AY399039	AY399039 Homo sapi
7	951.4	32.8	1887	9 AY399040	AY399040 Pan trogl
8	934.4	32.2	7004	3 CR748209	CR748209 Homo sapi
9	721	24.8	746	4 B1653092	B1653092 60330616
10	720	24.8	795	6 CA131312	CA131312 UI-M-FW0-
11	714	24.6	727	7 CF729223	CF729223 UI-M-FW0-
12	700.2	24.1	992	4 BG915267	BG915267 602815231
13	688	23.7	705	7 CF726196	CF726196 UI-M-FW0-
14	683.6	23.6	726	6 CB248251	CB248251 UI-M-FW0-
15	672	23.2	672	7 CO424322	CO424322 UI-M-FW0-
16	667.4	23.0	716	7 CO041415	CO041415 UI-M-FW0-
17	663	22.8	986	7 CF584809	CF584809 AGENCOURT
18	660.8	22.6	816	4 B1143828	B1143828 602907107
19	655.4	22.6	704	7 CO041178	CO041178 UI-M-FW0-
20	654	22.5	893	2 BF785290	BF785290 602108490
21	653.2	22.5	713	2 BB625283	BB625283 BB625283
22	652.4	22.5	655	7 CK781308	CK781308 UI-M-G10-
23	651.6	22.5	805	4 BG174288	BG174288 602334589
24	646	22.3	646	4 BG694881	BG694881 NISC_1v09

25	645.8	22.3	948	2 BF789245	BF789245 602105144
26	639.4	22.0	722	4 B1143493	B1143493 602907673
27	639	22.0	669	2 BB078728	BB078728 BB078728
28	627.6	21.6	782	7 CN261003	CN261003 1700004247
29	626	21.6	939	2 BE912458	BE912458 601666074
30	624.6	21.5	1069	5 BM927376	BM927376 AGENCOURT
31	615.6	21.2	1022	5 BQ062868	BQ062868 AGENCOURT
32	608	21.0	608	7 CK625537	CK625537 mJ10904.Y
33	600.8	20.7	616	4 B1647932	B1647932 UI-M-FW0-
34	599.4	20.7	660	7 CO043728	CO043728 UI-M-FW0-
35	596.2	20.5	1005	5 BQ060729	BQ060729 AGENCOURT
36	596.2	20.5	1029	5 BQ059508	BQ059508 AGENCOURT
37	596	20.5	674	5 BQ445704	BQ445704 UI-M-FW0-
38	593	20.4	600	4 BG806175	BG806175 2021-17 M
39	593	20.4	673	5 BQ746210	BQ746210 UI-M-FW0-
40	584.2	20.1	1012	5 BQ064231	BQ064231 AGENCOURT
41	582.8	20.1	818	7 CK483904	CK483904 AGENCOURT
42	562	19.4	1042	5 BM909401	BM909401 AGENCOURT
43	558.8	19.3	667	4 B1648260	B1648260 603378202
44	558.6	19.2	865	4 BG172224	BG172224 602333366
45	556	19.2	556	6 BY704937	BY704937 BY704937

ALIGNMENTS

RESULT 1	AK004737	2899 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK004737				
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013B22 product:weakly similar to PROABBL SERINE/THREONINE-PROTEIN KINASE KIA00537 (BC 2.7.1.-) [Homo sapiens], full insert sequence.				
ACCESSION	AK004737	GI:26334437			
VERSION	AK004737.2				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL	99279253				
MEDLINE	10349636				
PUBMED	10349636				
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, T., Nishi, K., Kizunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL	20530913				
MEDLINE	11076861				
PUBMED	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
JOURNAL	5				
REFERENCE					

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2899)
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hizmoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Katsukawa, T., Kato, H., Kawai, I., Kojima, Y., Komano, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okita, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toga, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-re@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT On Dec 10, 2002 this sequence version replaced gi:12836134. Please visit our web site (http://genome.gsc.riken.jp/) for further details.
FEATURES cDNA library was prepared and sequenced in Mouse Genome Encyclopædia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCGCGCACTCGAGTGTGTTTGTGTTTGTGTTT 3']. cDNA was prepared by using tritaholose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.
SOURCE location/Qualifiers
1. .2899
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:1200013B22"
/db_xref="taxon.10090"
/clone="1200013B22"
/sex="male"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
94. .1989
/note="unnamed protein product; putative weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|O60285, evidence: FAS2, 56.8%ID, 98.6%length, match=1849)"
/codon_start=1
/protein_id="BAB23518.1"
/db_xref="GI:12836135"
/translation="MESVALIQRPSQPSALASERAPLADGLIKSPYLMKKQAV KRRHKNILNRHYEFTLGKGYGKAKRSESGRLVAIKSRKDKIOEOLILNR RLIEIMSLNHPHIALHEVENSRSKIVIMEVYASRDLDYISERPSLRABHFF ROIVSLAHVCHONGIVARDIKLENIILLDANKIKAPDGSINLHKPKIQTGSGSL YASPELYNGKPYGPPEVDSMSGLYLTILVGHTEPFGQDHKTLVKISGAVREPK PSADCGIIRWLWNPFRTRATLEBVAHVMWVMWYTTGVEQELAREGSDPFGA SMADWRSRSPLEENGAKVCSFPGSGVSPOLAEQHSIKSKSRKMDAONLQ GDPAEDTSPRSGSLKPKGIKKSKSGVQEDPOELRIPTDPGQVPAVSLI PRGILKKSRORSRGVYSPSPSGSGSLASDVYFGDPEVDSKPSGAVSBDLSISESTQDL LKNGKSRFALGTTSTPGSGDQLASSPRAPRSPGAVSBDLSISESTQDL PERLPTPLRGCVSDNLRGLDEOPPSGLRMMQESIGDSFCSTLDCQEVTAAYRAL GTSKGS"

QY 991 TGAACCCACCCGTCGAGCCCACTGAGAGATGATAGCCAGTCAATTGGTGGTCAACTGG 1050
DB 962 TGAACCCACCCGTCGAGCCCACTGAGAGATGATAGCCAGTCAATTGGTGGTCAACTGG 1021
QY 1051 GTTACACACCCGAGTGGGAGAACAGAGAGCCCTGCTGAGGAGTGGGCACTTATGTGTG 1110
DB 1022 GTTACACACCCGAGTGGGAGAACAGAGAGCCCTGCTGAGGAGTGGGCACTTATGTGTG 1081
QY 1111 ACTTGGCCGGGCTCCATGGCGGAGTGTATAGTGTCTCTGCGGCCCTCTCTGAGAG 1170
DB 1082 ACTTGGCCGGGCTCCATGGCGGAGTGTATAGTGTCTCTCTGCGGCCCTCTCTGAGAG 1141
QY 1171 ATGAGCCCAAGGTGTGAGCTTCTTCAAGCAGACGTGCGCGGAGTGGAGCCTGTAC 1230
DB 1142 ATGAGCCCAAGGTGTGAGCTTCTTCAAGCAGACGTGCGCGGAGTGGAGCCTGTAC 1201
QY 1231 CTGGGCTGAGCGGCAATTCTTTAAGAGTCCCGAAGAGATGACATGGCTCAA 1290
DB 1202 CTGGGCTGAGCGGCAATTCTTTAAGAGTCCCGAAGAGATGACATGGCTCAA 1261
QY 1291 ATGCGAAGGTGACCGGCTGAGAGTACTCTCTCTGCGCGGAGAGAGCCTTAAGC 1350
DB 1262 ATGCGAAGGTGACCGGCTGAGAGTACTCTCTCTGCGCGGAGAGAGCCTTAAGC 1321
QY 1351 TTCCGAAAGGCAATTCTCAAGAAAAAGTCTCTACCTGCTCAGGAGGATACAGAGAGAC 1410
DB 1322 TTCCGAAAGGCAATTCTCAAGAAAAAGTCTCTACCTGCTCAGGAGGATACAGAGAGAC 1381
QY 1411 CTCAGAACTCAGACCGGTGCTGATATCTCAGGAGCAGCTGTCTGCTGTATCTCTGC 1470
DB 1382 CTCAGAACTCAGACCGGTGCTGATATCTCAGGAGCAGCTGTCTGCTGTATCTCTGC 1441
QY 1471 TCCCAAGAAAAAGATCTTTAAGAGTCTGACAGCGTGAATCTGGTATCTACTCTCTC 1530
DB 1442 TCCCAAGAAAAAGATCTTTAAGAGTCTGACAGCGTGAATCTGGTATCTACTCTCTC 1501
QY 1531 CAGAGCCCAAGGTGTGAGGAGCTCTTACAGCAGCAGTATGTTTGTGAGTGGAGAC 1590
DB 1502 CAGAGCCCAAGGTGTGAGGAGCTCTTACAGCAGCAGTATGTTTGTGAGTGGAGAC 1561
QY 1591 CCGTGAAGCAGAACTCTCAGAGCTTCAAGGCTCTCTCTCAGCCGAGAGGCAATTCTCA 1650
DB 1562 CCGTGAAGCAGAACTCTCAGAGCTTCAAGGCTCTCTCTCAGCCGAGAGGCAATTCTCA 1621
QY 1651 AACTCAATGGCAATTTCTCCGAGCAGCTTGAAGAGCACTACCCCTAGCACCCTTGGCT 1710
DB 1622 AACTCAATGGCAATTTCTCCGAGCAGCTTGAAGAGCACTACCCCTAGCACCCTTGGCT 1681
QY 1711 CCGTGAAGCAGAACTCTCAGAGCTTCAAGGCTCTCTCTCAGCCGAGAGGCTG 1770
DB 1682 CCGTGAAGCAGAACTCTCTCCATCTCTGAGCCCGGCGAGCCGCTCAAGAGGCTG 1741
QY 1771 TGAAGTGAAGCAGAGCTCTCTCTCAGAGTCTTGAACAATTGAAGCTTGAAGCTG 1830
DB 1742 TGAAGTGAAGCAGAGCTCTCTCTCAGAGTCTTGAACAATTGAAGCTTGAAGCTG 1801
QY 1831 TTCCGAAACCCCACTGAGGAGCTGTGTCTGTGACAACCTGAGGAGGCTTGAAGAGC 1890
DB 1802 TTCCGAAACCCCACTGAGGAGCTGTGTGTGTGACAACCTGAGGAGGCTTGAAGAGC 1861
QY 1891 CTCCCTCAAGAGGTGTAAGAGTGTGAGGAGGAGAACTCTGGGGAGATAGCTGCTTCTC 1950
DB 1862 CTCCCTCAAGAGGTGTAAGAGTGTGAGGAGGAGAACTCTGGGGAGATAGCTGCTTCTC 1921
QY 1951 TGAAGAGCTGCAAGAGGTGATCTGAGCTTACAGACAGCCCTAGAGATCTGCTCAAGC 2010
DB 1922 TGAAGAGCTGCAAGAGGTGATCTGAGCTTACAGACAGCCCTAGAGATCTGCTCAAGC 1981
QY 2011 TGAAGTGAAGAGGAGATGTGTCTTGAATGTGGGTGAGCTTGAAGAGGCTTGAAGAG 2070
DB 1982 TGAAGTGAAGAGGAGATGTGTCTTGAATGTGGGTGAGCTTGAAGAGGCTTGAAGAG 2041

QY 2071 GAACCTGGGCTGGATTTCTCCAGTGAATGAGTACATCAAGGGCTCTAGCTCTGACGCC 2130
DB 2042 GAACCTGGGCTGGATTTCTCCAGTGAATGAGTACATCAAGGGCTCTAGCTCTGACGCC 2101
QY 2131 TGAAGTGAAGAGTGAAGAGAAATGAGATGATGAGTGAAGAGAAATGAGAACTTGGC 2190
DB 2102 TGAAGTGAAGAGTGAAGAGAAATGAGATGATGAGTGAAGAGAAATGAGAACTTGGC 2161
QY 2191 TGCCGAGTGTATAGTGGGAGTGGCTGAAGAGTGCCTACCTCTTGTGCGCATAGTGTG 2250
DB 2162 TGCCGAGTGTATAGTGGGAGTGGCTGAAGAGTGCCTACCTCTTGTGCGCATAGTGTG 2221
QY 2251 ACCATGACATTTCCACCCCTGTTCTGTGCTGCACTTCAATAGTTCTGTTCCAT 2310
DB 2222 ACCATGACATTTCCACCCCTGTTCTGTGCTGCACTTCAATAGTTCTGTTCCAT 2281
QY 2311 CAACCAACAGGTTAAGAACCTCTGATCTCTGCGGAGTGAATGTGATGAGCAGCATTA 2370
DB 2282 CAACCAACAGGTTAAGAACCTCTGATCTCTGCGGAGTGAATGTGATGAGCAGCATTA 2341
QY 2371 TAGAGAGAAACAGCTCTGATTTCCATCTCTGCTGCTGCACTTCAATAGTTCTGTTCCAT 2430
DB 2342 TAGAGAGAAACAGCTCTGATTTCCATCTCTGCTGCTGCACTTCAATAGTTCTGTTCCAT 2401
QY 2431 GACTCGAGCCGCTGTTGACTTCACTCAAGAGGAGCAGATGCCCTGAGCCCATCTTA 2490
DB 2402 GACTCGAGCCGCTGTTGACTTCACTCAAGAGGAGCAGATGCCCTGAGCCCATCTTA 2461
QY 2491 GATCTCAGAGCTTGAACCTTGAAGCTGTTCTTGAATCCAGATGTGATGATGATGCTG 2550
DB 2462 GATCTCAGAGCTTGAACCTTGAAGCTGTTCTTGAATCCAGATGTGATGATGATGCTG 2521
QY 2551 TTTCTCAGGCAACAGGAGCCTAGATGAGTGAATTTATTTTGTGATCTCACT 2610
DB 2522 TTTCTCAGGCAACAGGAGCCTAGATGAGTGAATTTATTTTGTGATCTCACT 2581
QY 2611 TCTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGAAGTATTTGCTGCTT 2670
DB 2582 TCTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGAAGTATTTGCTGCTT 2641
QY 2671 CAATATGTGATGATGCTGTTCTGAGGAGCTCCAGCTGTGCACTGAAGTTATGATCAGA 2730
DB 2642 CAATATGTGATGATGCTGTTCTGAGGAGCTCCAGCTGTGCACTGAAGTTATGATCAGA 2701
QY 2731 GAAGTATTTGGCAATGATGCTCTATTCAGAGGAGGAGTGGGAGCTTTTCAATGAT 2790
DB 2702 GAAGTATTTGGCAATGATGCTCTATTCAGAGGAGGAGTGGGAGCTTTTCAATGAT 2761
QY 2791 GTCTTGAAGCAGCTGCTGATGATGATCTCAGATCCCTTCAACCCAGAGCTGGCACCCTC 2850
DB 2762 GTCTTGAAGCAGCTGCTGATGATGATGATCTCAGATCCCTTCAACCCAGAGCTGGCACCCTC 2821
QY 2851 CTTGATCTTCACTGTGGCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2899
DB 2822 CTTGATCTTCACTGTGGCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2870

RESULT 2
AK033672
LOCUS
DEFINITION
AK033672 3098 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130215K18 product:weakly similar to PROABAE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) [Homo sapiens], full insert sequence.
ACCESSION
AK033672
VERSION
AK033672.1 GI:26329364
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS
Carninci, P. and Hayashizaki, Y.

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 992799253 103496536 2
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 204993159 11042159 3
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Smit, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 5 (bases 1 to 3098) Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirokawa, I., Kanokawa, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akehira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyo, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, T.
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers 1. .3098 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM.DB:9130215K18" /db_xref="taxon:10090" /clone="9130215K18"
FEATURES	Source

QY 631 GCCACGAAAGGAGTGTTCACCGAGATCTCAAGCTGGAACAACTCTTGTAGATGCA 690
 Db 601 GCCACGAAAGGAGTGTTCACCGAGATCTCAAGCTGGAACAACTCTTGTAGATGCA 660
 QY 691 ATGGAACATCTAATATGCTGATCTTGGCCCTCTCCAACTGTATACCAAGAGATTC 750
 Db 661 ATGGAACATCTAATATGCTGATCTTGGCCCTCTCCAACTGTATACCAAGAGATTC 720
 QY 751 TCCAGAGTTCGTGGAGAGCCCTCTCTAGCCCTGAGTGTATGTAAGAGAGCCCT 810
 Db 721 TCCAGAGTTCGTGGAGAGCCCTCTCTAGCCCTGAGTGTATGTAAGAGAGCCCT 780
 QY 811 ATGTGGCCCAAGAGTGAAGCTGTGTCTGTGGGCTTCTCTGTATCACTCTGTATG 870
 Db 781 ATGTGGCCCAAGAGTGAAGCTGTGTCTGTGGGCTTCTCTGTATCACTCTGTATG 840
 QY 871 GCACCATGCTTTTGAAGGAGATCATTAACACTGTGTAGCAAACTCATACGGGG 930
 Db 841 GCACCATGCTTTTGAAGGAGATCATTAACACTGTGTAGCAAACTCATACGGGG 900
 QY 931 CTTAACGAGAGCCGCAAGCCGATCCGATGCTGTGGCTGTATCCGGGTGTATAGG 990
 Db 901 CTTAACGAGAGCCGCAAGCCGATCCGATGCTGTGGCTGTATCCGGGTGTATAGG 960
 QY 991 TGAACCCCAACCCGTCGGGCACTGTGAGATGTAGCCAGTCAATGTGTGTCAACTGG 1050
 Db 961 TGAACCCCAACCCGTCGGGCACTGTGAGATGTAGCCAGTCAATGTGTGTCAACTGG 1020
 QY 1051 GTTACACCAACCGAGTGGGGACAAGAAAGCTGTGGTGGAGGCTTATAGTGTG 1110
 Db 1021 GTTACACCAACCGAGTGGGGACAAGAAAGCTGTGGTGGAGGCTTATAGTGTG 1080
 QY 1111 ACTTGGCCGGGCTCTCATGGCCGATGTGTATGCTGTCTCTGGCCCTCTCTGAGA 1170
 Db 1081 ACTTGGCCGGGCTCTCATGGCCGATGTGTATGCTGTCTCTGGCCCTCTCTGAGA 1140
 QY 1171 ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGTGCGGGAGTGAAGCACTGTAC 1230
 Db 1141 ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGTGCGGGAGTGAAGCACTGTAC 1200
 QY 1231 CTGGGCTGAGAGCGGCAACTTCTTTAAGAAAGTCCGAAAGAGATGATGAGCTCAA 1290
 Db 1201 CTGGGCTGAGAGCGGCAACTTCTTTAAGAAAGTCCGAAAGAGATGATGAGCTCAA 1260
 QY 1291 ATCTGCAAGTGAACCGGCTGAGATCTCTTCTCGCCCTGGCAAGAGCACTTAAAC 1350
 Db 1261 ATCTGCAAGTGAACCGGCTGAGATCTCTTCTCGCCCTGGCAAGAGCACTTAAAC 1320
 QY 1351 TTCCGAAAGGCAATCTCAAGAAAGTCTCTACCTCGTCAAGGAGGATACAGAGAGCC 1410
 Db 1321 TTCCGAAAGGCAATCTCAAGAAAGTCTCTACCTCGTCAAGGAGGATACAGAGAGCC 1380
 QY 1411 CTGAGAACTCAGACCGGTGCTGTATCTCAAGGAGCACTGTCCCTGTATCTCTGC 1470
 Db 1381 CTGAGAACTCAGACCGGTGCTGTATCTCAAGGAGCACTGTCCCTGTATCTCTGC 1440
 QY 1471 TCCCAAGAAAGGCAATCTTAAAGATCTGACAGCGTGAATCTGTATCTACTCTCTC 1530
 Db 1441 TCCCAAGAAAGGCAATCTTAAAGATCTGACAGCGTGAATCTGTATCTACTCTCTC 1500
 QY 1531 CAGAGCCCAAGGATCTGGGGACCTTGAAGCGCAAGTGTGTGTGTGTGTGTGTGTGT 1590
 Db 1501 CAGAGCCCAAGGATCTGGGGACCTTGAAGCGCAAGTGTGTGTGTGTGTGTGTGTGT 1560
 QY 1591 CCGTGGAGCAGAGTCTTCACAGGCTTCAAGGCTCTCTCTCACCGCAAGGAGCAATTC 1650
 Db 1561 CCGTGGAGCAGAGTCTTCACAGGCTTCAAGGCTCTCTCTCACCGCAAGGAGCAATTC 1620
 QY 1651 AACTCAATGGAAGTTCCTCCGCAAGCTTAAAGAGCACTAACCCCTTGAAGCTTGGCT 1710
 Db 1621 AACTCAATGGAAGTTCCTCCGCAAGCTTAAAGAGCACTAACCCCTTGAAGCTTGGCT 1680

QY 1711 CCTTGAACCAATGAGGCTCTCTCCATCTCTGAGAGCCGCGCCCAAGCCCTCTCAAGGGCTG 1770
 Db 1681 CCTTGAACCAATGAGGCTCTCTCCATCTCTGAGAGCCGCGCCCAAGCCCTCTCAAGGGCTG 1740
 QY 1771 TGAATGAGGACAGCATCTGTGCTCTCGAGTCTTTTGAACCAATTTGAGCTTGTGAGAGCTC 1830
 Db 1741 TGAATGAGGACAGCATCTGTGCTCTCGAGTCTTTTGAACCAATTTGAGCTTGTGAGAGCTC 1800
 QY 1831 TTCCGAAACCCCACTGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1890
 Db 1801 TTCCGAAACCCCACTGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
 QY 1891 CTCCCTCAAGAAAGTCTGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1950
 Db 1861 CTCCCTCAAGAAAGTCTGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
 QY 1951 TGAAGACTGCAAGAGTGAATGAGCTGACCTTACCAAGAGCCCTTGAAGTGTGTGTGTGT 2010
 Db 1921 TGAAGACTGCAAGAGTGAATGAGCTGACCTTACCAAGAGCCCTTGAAGTGTGTGTGTGT 1980
 QY 2011 TCAGCTGAGGAAAGGAGATGT 2070
 Db 1981 TCAGCTGAGGAAAGGAGATGT 2040
 QY 2071 GAACCTGTGGTGGATTCCTTCAAGTGAATAGATGATCAATCAAGGCTCTTACGCTGTGACGC 2130
 Db 2041 GAACCTGTGGTGGATTCCTTCAAGTGAATAGATGATGATCAATCAAGGCTCTTACGCTGTGACGC 2100
 QY 2131 TGAATGAACCTGAAGATGAGAGAAATGCAATTTATGTGTGAAGAATGGAACCTTTGC 2190
 Db 2101 TGAATGAACCTGAAGATGAGAGAAATGCAATTTATGTGTGAAGAATGGAACCTTTGC 2160
 QY 2191 TGCCGAGGTGTATAGTGGGGTGGCTGAAGGTGCTTACCTCTTGTGGCATGAGTGTG 2250
 Db 2161 TGCCGAGGTGTATAGTGGGGTGGCTGAAGGTGCTTACCTCTTGTGGCATGAGTGTG 2220
 QY 2251 ACCATGACATTTTCCACCCCTGTCTGTGCTGACCTTCAATAGTTTCTGTTCAT 2310
 Db 2221 ACCATGACATTTTCCACCCCTGTCTGTGCTGACCTTCAATAGTTTCTGTTCAT 2280
 QY 2311 CAACCAACAGGTTTGAACCTTGAATCTCTGTGGAGTGAATGTGTATGTATGTATGTAT 2370
 Db 2281 CAACCAACAGGTTTGAACCTTGAATCTCTGTGGAGTGAATGTGTATGTATGTATGTAT 2340
 QY 2371 TAGAGAGAAACACCTCTGTGTTCCATCTCTGTGCTGTGATCTCTCAAGACCTGGGAA 2430
 Db 2341 TAGAGAGAAACACCTCTGTGTTCCATCTCTGTGCTGTGATCTCTCAAGACCTGGGAA 2400
 QY 2431 GACTCGAACCGCTGTGTGATCTTCAATCTCAAGGAGCAAGATGCCCTGTGACCCCATCTTA 2490
 Db 2401 GACTCGAACCGCTGTGTGATCTTCAATCTCAAGGAGCAAGATGCCCTGTGACCCCATCTTA 2460
 QY 2491 GATCTCAGAGCTTGAACCTTGAAGCTGTCTTATGATCCAGATGTGTATGTATGTATGT 2550
 Db 2461 GATCTCAGAGCTTGAACCTTGAAGCTGTCTTATGATCCAGATGTGTATGTATGTATGT 2520
 QY 2551 TTTCTCAGGCAAGGAGACCTAGATGTGCTGATCTTATTTATTTTGTGTGTGTGTGTGT 2610
 Db 2521 TTTCTCAGGCAAGGAGACCTAGATGTGCTGATCTTATTTATTTTGTGTGTGTGTGTGT 2580
 QY 2611 TCTGT 2670
 Db 2581 TCTGT 2640
 QY 2671 CAATTAATGTATGATCTGT 2730
 Db 2641 CAATTAATGTATGATCTGT 2700
 QY 2731 GAAGTATTTGGCAATGATGCTCTCTTATTTCAAGGGGGTGTGGGGCTTTTCAATGAT 2790
 Db 2701 GAAGTATTTGGCAATGATGCTCTCTTATTTCAAGGGGGTGTGGGGCTTTTCAATGAT 2760
 QY 2791 GTCTTGAACATGTCTGTGATGTGATCTCAAGTCTCTTCAACCAAGGCTGTGCAACCTC 2850

Db	2761	GCTTGACCACTGTCGTGGATTCCTCAGTCCCTTCACACCAGGCTG3CACCTGC	2820
Oy	2851	CCTCATCTTCATCTGTGGCCAAACAAAAAAAAAAAAAAAAAAAAA	2899
Db	2821	CCTCATCTTCATCTGTGGCCAAACAAAAAAAAAAAAAAAAAAAAACA	2869
RESULT 3			
AKO34082			
LOCUS		2869 bp mRNA linear HTC 03-APR-2004	
DEFINITION	Mus musculus adult male dienecephalon cDNA, RIKEN full-length		
	enriched library, clone:9330154N4 product:weakly similar to		
	PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)		
	[Homo sapiens], full insert sequence.		
ACCESSION	AKO34082		
VERSION	AKO34082.1	GI:26329672	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Weich. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			

TITLE	COMMENT	FEATURES
JOURNAL	Direct Submission Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suohi-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research@riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	source
	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers	1..2869 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM.DB:9330154N24" /db_xref="taxon.10090" /clone="9330154N24" /sex="male" /tissue_type="diencephalon" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 98..2017 /note="unnamed protein product; putative weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT O60285, evidence: PASTY, 56.8%ID, 98.6%length, match=1849)" /codon_start=1 /protein_id="BAC8575.1" /db_xref="gi:26329673" /translation="MESGALLQRPSPAPASALASBSARPIADGLISPYRLMKQAVKRHHKNLRHYEFLBTGKGTGYAKKRESGRVAIKSIRKDKIKBOEDLHRIKREIMSLNHPHIAIHVEGRSLVAVPENSKIIVIMEYASRGLADYISEPRRLSEBRDHPFPROISALHYCHONGIVHRLKLENTILINDPGLSNLYHKGFLEOTPCGSPVSPASPEIYVNGKPYGPEVDSWLSVILYLVHGMPPDGDHKLTVQINS GAYEPPKPSDACGLIRMLWMVNTFRATLIEDVASHMMVMNGITTVGEOLAEHGGH PSQNGRASMADWLRRSSRPLLENGAVCSFPQHVGGSTVGLRQSLKSKREK NDMAMLOQDGPABDTSSRPKGSILKPKGLIKKSSSTSGEVDQDLRPPVPCGSPV PVPVPLPKRGLIKKRSRSGSYSPESPESEGLDASPVFGSPVBEKSPVBEKSTLSS ILHRRGLIKNGKESRSTALEGTTSPFSGSLDOLASHPARPSPGASPESTLSS ESFQDLPRLPRTPIRGCYSVNLNGLBEQPEBGLKRWQESLSDSCSTLDCQERY TAAVRQALGICSKLS"
	Query Match 96.6%; Score 2804.4; DB 3; Length 2869; Best Local Similarity 99.1%; Pred. No. 0; Matches 2839; Conservative 0; Mismatches 1; Indels 24; Gaps 1;	
ORIGIN		
Query Match		
Best Local Similarity		
Matches 2839;		
Conservative		
Mismatches		
Indels		
Gaps		
1;		
31		
6		
91		
66		
151		
126		
211		
186		
271		
246		
305		

Db	2466	CAGATGCCCTGGAGCCCATCTTGAATCTCAGAGACTTGAACCTTGAAGCTGTTCTAAGT	2525		
Qy	2527	ACCCAGATGTGGATGATGATGCTCTGTTTCTCAGGCCACGAGACCTAGAAATGTGCTGACTT	2586		
Db	2526	ACCCAGATGTGGATGATGATGCTCTGTTTCTCAGGCCACGAGACCTAGAAATGTGCTGACTT	2585		
Qy	2587	ATTATATTTTTTTTGGATCTCTACATTCCTGTTTTTGGTTTTTGGTTTTGTTGTTGTTTTTGG	2646		
Db	2586	ATTATATTTTTTTTGGATCTCTACATTCCTGTTTTTGGTTTTTGGTTTTGTTGTTGTTTTTGG	2645		
Qy	2647	TTTTTAAGTGAATTTTGTGCTTTCAATATATGTGAATGCTGTGTTCTGGGAACTCCACT	2706		
Db	2646	TTTTTAAGTGAATTTTGTGCTTTCAATATATGTGAATGCTGTGTTCTGGGAACTCCACT	2705		
Qy	2707	GTGCCACTGAAGTTATGTATGACAGAGAAATTTGGCAATGATATGTCCTCTATTCAGAGGG	2766		
Db	2706	GTGCCACTGAAGTTATGTATGACAGAGAAATTTGGCAATGATATGTCCTCTATTCAGAGGG	2765		
Qy	2767	GGTGGGGGGCGTTTTCATATGTATGTCTTGAAGCACTGTCTGATTTGAATCTCCAGTCCCT	2826		
Db	2766	GGTGGGGGGCGTTTTCATATGTATGTCTTGAAGCACTGTCTGATTTGAATCTCCAGTCCCT	2825		
Qy	2827	TCACACCCCAAGAGCTGGCCACCCTCCATCTTCAATCATCTGTGGCC	2870		
Db	2826	TCACACCCCAAGAGCTGGCCACCCTCCATCTTCAATCATCTGTGGCC	2869		
RESULT 4	AY399041	1896 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY399041	Mus musculus HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence,			
ACCESSION	AY399041	genomic survey sequence.			
VERSION	AY399041.1	GI:39755030			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.C., Adams,M.D. and Cargill,M.				
TITLE	Interfering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PIUMED	14671302				
REFERENCE	2 (bases 1 to 1896)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.C., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	location/Qualifiers				
source	1..1896				
gene	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	<1..>1896				
	/locus_tag="HCM0091"				
ORIGIN					
Query Match	61.2%;	Score 1776;	DB 9;	Length 1896;	
Best Local Similarity	93.7%;	Pred. No. 0;			
Matches 1776;	Conservative 0;	Mismatches 120;	Indels 0;	Gaps 0;	
123	ATGAGATCGGTGGCTTACTCTCAGAGCCGACGACGAGCTCCCTGGGCTTCGACCTGGCC	182			

D	b	1	ATGGAAGTCGGAGGCGCTTAACTTCACAGCGCCGAGCCAGGAGCTCCTCTGAGCTTCGCGCTTGACC	60
O	y	183	TCGGAGAGCGCCCGCGCGCTGCGCGAGCGGGCTCATCAAGTCGACTTAAACCTTGATGAAG	242
D	b	61	TCGGAGAGCGCCCGCGCGCTGCGCGAGCGGGCTCATCAAGTCGACTTAAACCTTGATGAAG	120
O	y	243	AAGCAGCGCGGTGAAGCGGCACCATCACAAACAACTCGCGGCACCGCTACGAGTTCCTG	302
D	b	121	AAGCAGCGCGGTGAAGCGGCACCATCACAAACAACTCGCGGCACCGCTACGAGTTCCTG	180
O	y	303	GAGAGCGCTGGGSCAAGGGCACCTACGGGAAGGTGAAGAAGCACAGAGAGCTCGGGGCGT	362
D	b	181	GAGAGCGCTGGGSCAAGGGCACCTACGGGAAGGTGAAGAAGCACAGAGAGCTCGGGGCGT	240
O	y	363	CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAAGTGAAGAGATCTTGCTGCAC	422
D	b	241	CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAAGTGAAGAGATCTTGCTGCAC	300
O	y	423	ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCACATCATTTGCCATCAT	482
D	b	301	ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCACATCATTTGCCATCAT	360
O	y	483	GAACTGTTTGAGATACAGCAAGAAATTGTATTTGTATGAGATGTGCAAGCCGAGGGGAT	542
D	b	361	GAACTGTTTGAGATACAGCAAGAAATTGTATTTGTATGAGATGTGCAAGCCGAGGGGAT	420
O	y	543	CTGTATGATTTCATCAAGTGAAGCGGCACACGGCTGAGTGAAGCGGCACGACGATTTCTTC	602
D	b	421	CTGTATGATTTCATCAAGTGAAGCGGCACACGGCTGAGTGAAGCGGCACGACGATTTCTTC	480
O	y	603	CGACAGATCGTGTCTGCGCTTGCACTACTGCCACCAAGCGGATCGTTTCAACGAGATCTC	662
D	b	481	CGACAGATCGTGTCTGCGCTTGCACTACTGCCACCAAGCGGATCGTTTCAACGAGATCTC	540
O	y	663	AAGCTGGAAGAAACATCCCTTTCTAGATGCCAATGGAACATCAAAGATTGCTGACTTTGGGCTC	722
D	b	541	AAGCTGGAAGAAACATCCCTTTCTAGATGCCAATGGAACATCAAAGATTGCTGACTTTGGGCTC	600
O	y	723	TCCAACCTGTACCAAAAGCAAGTCTCTCACAAGCTTCTGTGGAGACCCCTCTTAGGCC	782
D	b	601	NN	660
O	y	783	TCGCGTAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGGTGAACAGCTGTCTCTG	842
D	b	661	NN	720
O	y	843	GGCGTTCTCCCTGTATCATCTCTGTGTGATGAGCAACATGCGCTTTTGAACGGGCAAGATCATATAA	902
D	b	721	GGCGTTCTCCCTGTATCATCTCTGTGTGATGAGCAACATGCGCTTTTGAACGGGCAAGATCATATAA	780
O	y	903	ACACTGTGGAAGCAAAATCAAGTAACGGGGCTTAAACGTGAGCGGCCCAAGCGTCCGATGCC	962
D	b	781	ACACTGTGGAAGCAAAATCAAGTAACGGGGCTTAAACGTGAGCGGCCCAAGCGTCCGATGCC	840
O	y	963	TGTGGCGTGATCCGGTGGCTGTTAATGTGAACCCACCCGTGCGGCCACAATTGGAAGAT	1022
D	b	841	TGTGGCGTGATCCGGTGGCTGTTAATGTGAACCCACCCGTGCGGCCACAATTGGAAGAT	900
O	y	1023	GTAGCCAGTCAATTGTGGGTTCACCTGGGGTTTACCAACCGGAGTCCGGGGAACAGGAAGCC	1082
D	b	901	GTAGCCAGTCAATTGTGGGTTCACCTGGGGTTTACCAACCGGAGTCCGGGGAACAGGAAGCC	960
O	y	1083	CTGGGTGAAGGGTGGGACCCGTAAGTGGTCACTTTGGCGGGCGTCCATGGGGGATGTGGTTA	1142
D	b	961	CTGGGTGAAGGGTGGGACCCGTAAGTGGTCACTTTGGCGGGCGTCCATGGGGGATGTGGTTA	1020
O	y	1143	CGTGCCTCTTCGCGCCCTCTCTGGAATAAGAGCCAAAGGTGTGCACTTCTTCAACAG	1202
D	b	1021	CGTGCCTCTTCGCGCCCTCTCTGGAATAAGAGCCAAAGGTGTGCACTTCTTCAACAG	1080
O	y	1203	CAGGTGCCGGAGGTGAAGCATGTACTGTGGCTGGAGCGGCAACTTCTTTAAGAG	1262

Db 1081 CAGCTGCCGGAGGTGAGACATCTGACCTGAGCGGCAACATTTCTTAAGAG 1140

Qy 1263 TCCGGAAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTCT 1322

Db 1141 TCCGGAAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTCT 1200

Qy 1323 TCTGCGCTGGCAAGAGACCTTAAGCTTCCGAAAGGCAATTCAGAAAAAGTCTCT 1382

Db 1201 TCTGCGCTGGCAAGAGACCTTAAGCTTCCGAAAGGCAATTCAGAAAAAGTCTCT 1260

Qy 1383 ACCCTGTCAGGGGAGGTACAGAGACCTTACAGAACTCAACCGGTCCTGATACCTCA 1442

Db 1261 ACCCTGTCAGGGGAGGTACAGAGACCTTACAGAACTCAACCGGTCCTGATACCTCA 1320

Qy 1443 GGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502

Db 1321 GGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

Qy 1503 CAGCGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562

Db 1381 CAGCGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

Qy 1563 GCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622

Db 1441 GCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

Qy 1623 CTCTCTCTCCAGCGGAGGAGCTTCTCAAACTCAATGGCAAGTTCCTCCGACAGCTTA 1682

Db 1501 CTCTCTCTCCAGCGGAGGAGCTTCTCAAACTCAATGGCAAGTTCCTCCGACAGCTTA 1560

Qy 1683 GAAGGCACTACCTCCAGGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742

Db 1561 GAAGGCACTACCTCCAGGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

Qy 1743 GCCCGCCGAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802

Db 1621 GCCCGCCGAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

Qy 1803 TTGACCAATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1862

Db 1681 TTGACCAATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

Qy 1863 GTGAGCAACCTGAGGGGCTTGAAGAGCTTCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1922

Db 1741 GTGAGCAACCTGAGGGGCTTGAAGAGCTTCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1800

Qy 1923 GAATCTCTGGGGAGTACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982

Db 1801 GAATCTCTGGGGAGTACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

Qy 1983 AGACAAGCCCTAGGAATCTGCTCAAAAGCTGAGCTGA 2018

Db 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTGAGCTGA 1896

RESULT 5
 HSM801859
 LOCUS HSM801859 3443 bp mRNA linear HTC 22-SEP-2004
 DEFINITION Homo sapiens mRNA; cDNA DKFp434J037 (from clone DKFp434J037).
 ACCESSION AL136891
 VERSION AL136891.1 GI:12053280
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3443)
 Koehrer, K., Beyer, A., Nemes, H. W., Weill, B., Amdt, C., Oesanger, A.,
 Fodor, G., Han, M. and Wiemann, S.
 The German cDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuberberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by BMRP (Biomedical Research Center) at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZ434J037) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZ434J037
 Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES
 source
 1..3443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZ434J037"
 /db_xref="taxon:9606"
 /clone="DKFZ434J037"
 /issue_type="testis"
 /clone_id="434 (synonym: hte33). Vector pSport1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /note="hypothetical protein"
 1..3443
 /gene="DKFZ434J037"
 128..2014
 /gene="DKFZ434J037"
 /product="hypothetical protein"
 /protein_id="CAB6825.1"
 /db_xref="GO:09093"
 /db_xref="UniProt/TREMBL:Q9H093"
 /translation="MESLIVFARBSGPSPASAEILAPLAEGILKSPKLMKQAVRHH
 HKHLRIEPLLETLGKTKGKAKAESRLVATISIRDKIKEDDLHTRREIT
 IMSLNPHIILHREVPENSKIVIVNEVRSGLUYIISRQQLSEREASHFRQIV
 SAVHYQHONVVRHDLKENILIDANENIKIADGLENLHQGKFLQFCSPPLAASP
 EIVNGKEPTGPEVDSMSLGVLLYLIVNGVTFDGHDKILVQKLSNGAYRPPPSDA
 CGLIRMLMNPTRATLEDAVSHWNVGATRGVDEAPHEGSGPSDASADMA
 MRRSSRPLENGAKVCSFPKQARPGGSGTTPGLERHSLKSRKENDMAQSLSDTA
 DDIARHPKSNLKLPGKILKKVSASABGVQEDPEPSPITPASGQAPLIPKGIK
 KPRQRESGYVSPSPSSGSLDAGDVFSQDPEKQRPQASGLLHKKGLIKLNGKF
 SOTALBLAATTFGSLDELAPRPLAPASPSGAVSDSILSSSPFDLDPERLPEP
 PLRGCVSDNLTGLEBPSPSGPSCLRRMRDPLGDSCLFSLDCQEVATATYQALRYC
 SKLT"

ORIGIN
 Query Match 47.4%; Score 1374.2; DB 3; Length 3443;
 Best Local Similarity 81.6%; Pred. No. 7.5e-313;
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

Qy 82 CTACTGCTGCCGAGCCCACTGACCTGCGGCTCCCGGACATGAGTGGCTTAC 141

Db 87 CTACTGATTCCTCCGCGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146

Qy 142 TCCAGCGCCGAGCGAGGCTCCCTGCGGCTCCGCGCTGCGGCTGCGAGAGCGCCGCGCGC 201

Db 147 CGCGGCGCTCCGCGCCCACTCTCGGCGGAGAGCTA-----GCCGCGCGC 194

Qy 202 TGCGGAGCGGCTCATGAGTGGCTAACTCTGATGATGAGAGAGCGGCTGAGAGCGCGC 261

Db 195 TGCGGAGAGGCTATGAGTGGCGCAAGCCCTTAATGATAAGACAGCGGCTGAGAGCGCGC 254

Qy 262 ACATCAACAACAACCTGCGGACCGGCTTGAAGTCTTGAAGACGCTGGGCAAGGCGCA 321

Db 255 ACCACCAACAACAACCTGCGGACCGGCTTGAAGTCTTGAAGACGCTGGGCAAGGCGCA 314

Qy 322 CCTACGGGAGAGTGAAGAGGACGAGAGAGCTGCGGCGCTGCTGAGCCATCAAGTCA 381

Db 315 CTTACGGGAGAGTGAAGAGGCGCGGAGAGCTGCGGCGCTGCTGAGCCATCAAGTCA 374

Qy 382 TCAGAAAGACAAATCAAAATGAGAGAGATCTGCTGACATACGAGGAGATTTAGA 441

Db 375 TCCGGAAGACAAATCAGAGATGAGCAAGATCTGATGACATACGAGGGAGATTGGA 434
Qy 442 TCATGTTCTTCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAATGCA 501
Db 435 TCATGTCATCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAATGCA 494
Qy 502 GCAAGATTGATATGTCATGAGATATGACCGAGCGGATCTGATGATTAATCACTAGT 561
Db 495 GCAGATCTGATATGTCATGAGATATGACCGAGCGGATCTGATGATTAATCACTAGT 554
Qy 562 AGCGGCAACGAGTATGAGCGGAGCGGAGGATTTCTTCCGACAGATCTGTCGCC 621
Db 555 AGCGGCAACGAGTATGAGCGGAGCGGAGGATTTCTTCCGAGATTCCTCTGCG 614
Qy 622 TGCACTATGTCACCAAGACGGATTCCTTACCGAGATCTAGCTGAGAAATCTCTT 681
Db 615 TGCACTATGTCACCAAGATTCCTTACCGAGATCTAGCTGAGAAATCTCTT 674
Qy 682 TAGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 741
Db 675 TGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 734
Qy 742 GCAAGTCTTCAACGCTTCTGAGGAGCCCTCTCTACGCTCGCTGATGATGCAAG 801
Db 735 GCAAGTCTTCAACGATTCCTGAGGAGCCCTCTTATGCTCGCAGATGATGCAAG 794
Qy 802 GGAAGCCCTATATGAGGAGGAGGATGACAGCTGATGCTGAGGCTTCTCTTATCCTC 861
Db 795 GGAAGCCCTATCAACGAGGAGGATGACAGCTGATGCTGAGGCTTCTCTTATCCTC 854
Qy 862 TGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 921
Db 855 TGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 914
Qy 922 GTAAAGGAGCTTACCGTGAAGCGCCCAAGCGTCCATGCTGAGCTGATCCGATG 981
Db 915 GCAAGGAGCTTACCGGAGGAGGATGACAGCTGATGCTGAGGCTTCTCTTATCCTC 974
Qy 982 TGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 1041
Db 975 TGATGCAATGAGAAATCAAGATTGCTGCTTTGGCTCTCCAACTGATCAACAAG 1034
Qy 1042 TCAACTGAGGAGTACCAACGAGGAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1101
Db 1035 TCAACTGAGGAGTACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1094
Qy 1102 CTAGTGTGATCTTGGCGGAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1161
Db 1095 CTAGTGTGATCTTGGCGGAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1154
Qy 1162 TCCGGAAGATGAGACCAAGGATGAGCTTCTTCAAGACAGAGCTGCGGAGGAGGAG 1221
Db 1155 TCCGGAAGATGAGACCAAGGATGAGCTTCTTCAAGACAGAGCTGCGGAGGAGGAG 1214
Qy 1222 GCACTGATACCTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1281
Db 1215 GCACTGATACCTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1274
Qy 1282 TGGCTCAAAATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1341
Db 1275 TGGCTCAAAATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1334
Qy 1342 GCTTGAAGCTTCCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1401
Db 1335 ACCTCAAGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1394
Qy 1402 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1461
Db 1395 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1448
Qy 1462 TATCCCTGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1521
Db 1449 ---CGCTGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1505

Qy 1522 ACTCTCTCCAGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1581
Db 1506 ACTCTCTCCAGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1565
Qy 1582 GTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1641
Db 1566 GTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1625
Qy 1642 GCATTTCAATCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1701
Db 1626 GCATTTCAATCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1685
Qy 1702 CTTTGGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1761
Db 1686 CTTTGGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1745
Qy 1762 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1821
Db 1746 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1805
Qy 1822 CTGAAGCTCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1881
Db 1806 CTGAAGCTCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1865
Qy 1882 TTGAGAGGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1929
Db 1866 TTGAGAGGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1925
Qy 1930 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1989
Db 1926 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1985
Qy 1990 CCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2049
Db 1986 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2044
Qy 2050 GCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2092
Db 2045 GCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2087

RESULT 6
AY399039 1887 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION
Homo sapiens HCM0091 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY399039
VERSION
AY399039.1 GI:39755028
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

genomic survey sequence.
AY399040
VERSION AY399040.1 GI:3975029
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
AUTHORS 1 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1887)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submisison
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1887
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1887
/locus_tag="HCM0091"
ORIGIN
gene
Query Match 32.8%; Score 951.4; DB: 9; Length 1887;
Best Local Similarity 60.1%; Pred. No. 4.2e-213;
Matches 1119; Conservative 0; Mismatches 722; Indels 21; Gaps 2;
QY 169 CCTCCGCGCTGAGCGCTGCGAGAGCGCGCGCTGCGGAGCGGGCTCATCAAGTCGCGCTA 228
DB 35 CCACTCCCTCGGCGCGAGAGCTAGCCCGCGCTGCGGAGCGGGCTCATCAAGTCGCGCA 94
QY 229 AACCTGTATGAAGACGAGCGGTGAAGCGGCACATCAACAACACACTGCGGCA 288
DB 95 AGCCCTTAATGAAGAACGAGCGGTGAAGCGGCACACACACACACTGCGGCA 154
QY 289 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTACGGGAGGTGAAGAGCGAG 348
DB 155 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTACGGGAGGTGAAGAGCGAG 214
QY 349 AGAGCTCGGGGCGCTCTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 408
DB 215 AGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 274
QY 409 AGAGTCTGTGCAATACGAGGAGATGAGATCATGTCTTCACTCAACCCGCCACA 468
DB 275 AAGATCTGATGCAATACGAGGAGATGAGATCATGTCTTCACTCAACCCGCCACA 334
QY 469 TCAATGTCATTCATGAGATGTTTGAAGATAGCAGCAAGATTGTCTCATGAGATG 528
DB 335 TCAATGTCATTCATGAGATGTTTGAAGATAGCAGCAAGATTGTCTCATGAGATG 394
QY 529 CCAAGCGAGCGATCTGTATGATTAATCATCAGTAGCGGCCACGGCTGAGTAGCGGAGC 588
DB 395 NNN 454
QY 589 CCAAGCATTTCTCCGACAGATCGTGTCTGCCCTGCACTACTGCCACGAGACGGATCG 648
DB 455 NNN 514
QY 649 TTCAACGAGATCTCAAGCTGGAACAATCTCTTAGATGCCAATGAGAAATCAAGATTG 708

DB 515 TCACCGAGATCTCAAGCTGAGAACATCTCTTGTGATGCCAATGAGAAATCAAGANN 574
QY 709 CTGACTTTGGCTCTCCAACTGTACCAAAAGCAAGTTCTTCAGACGTTCTGTGAGA 768
DB 575 NNN 634
QY 769 GCCCTCTCTACGCTCGCTGTAGATAGTCAACGGAAGCCCTATGTGGGCCAGAGTGG 828
DB 635 NNN 694
QY 829 ACAGCTGTCTGTGGGGTCTCTCTGTACATCTCTGTGATGAGACCATGCTTTGACG 888
DB 695 NNN 754
QY 889 GGCAGATCATAAACAAGTGAAGCAAAATCAAGTACGAGGCTTACCTGAGCGGCCA 948
DB 755 NNN 814
QY 949 AGCCGTCCATGCTGTGGCTGTATCCGTGGCTTTAATGTGAACCCACCCGTGGG 1008
DB 815 NNN 874
QY 1009 CCAACTGAGAGATGTAGCAATGTAGTGTGCTGCACTGGGGTTACACACCGAGCTCG 1068
DB 875 CCACTGAGAGATGTAGCAATGTAGTGTGCTGCACTGGGGTTACACACCGAGCTCG 934
QY 1069 GGGAAACAGAAAGCTGCTGAGAGGTGGAGCACTTAAGTGTGACTTTGGCGGCTTCA 1128
DB 935 GAGAGCAGAGAGGCTCGGCAATGAGGTGGAGCACTTGCATGACTGTCTCGGCTCG 994
QY 1129 TGGCGAGCTGTGATAGTGTCTCTCTCGGCGCTCTCTGAGAAATGAGCAAGTGTGCA 1188
DB 995 TGGCTGACTGTGCTCGGCTGTCTCTCGGCGCTCTCTGAGAAATGAGGCAAGTGTGCA 1054
QY 1189 GCTTCTTCAAGACAGACGTGCGGAGGTGAGAACATGATCTGTGGCTGAGAGGCAAC 1248
DB 1055 GCTTCTTCAAGACAGACGTGCGGAGGTGAGAACACACCTGTGAGAGGCGCAGC 1114
QY 1249 ATTCTCTTAAGAAGTCCGAAAGAGATGATGATGCTCAAAATCTCAAGTGAACCCG 1308
DB 1115 ATTCTCTTAAGAAGTCCGAAAGAGATGATGATGCTCAAAATCTCAAGTGAACCCG 1174
QY 1309 CTGAGATACCTCTCTCTGCGCTTGGCAAGAGACCTTAAGTTCGAAAGGCAATTCTCA 1368
DB 1175 CTGATGACACTGCGCATGCGCTTGGCAAGAGACCTTAAGTTCGAAAGGCAATTCTCA 1234
QY 1369 AGAAAGTCTCTCTCTCTCTGCGCTTGGCAAGAGACCTTAAGTTCGAAAGGCAATTCTCA 1428
DB 1235 AGAAAGTCTCTCTCTCTCTGCGCTTGGCAAGAGACCTTAAGTTCGAAAGGCAATTCTCA 1294
QY 1429 TGCTGTATCTCAAGGAGCACTGTCTCTGTATCTCTGCTCCCAAGAAAGGCAATCC 1488
DB 1295 TCCTGTGAGACCTCAAGGAGCACTGTCTCTGTATCTCTGCTCCCAAGAAAGGCAATCC 1345
QY 1489 TTAAGAGTCTGACAGCGTGAATCTGTGTTACTTCTCTGAGAGCCGAGAGTCTG 1548
DB 1346 TCAAGAGCCCAACAGCGGAGGTCTGACTACTCTCTCCGAGCCGAGTGAATCTG 1405
QY 1549 GGGAACTCTTAAGACGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1608
DB 1406 GGGAGCTCANNAGAGCAGCGGAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1465
QY 1609 CACAGGCTTCAAGGCTCTCTCTCAACGCAAGGCAATTCACAATCAATGAGCAATTC 1668
DB 1466 NNN 1525
QY 1669 CCGCAGAGCTTAGAAGGCACTTACCTGAGCACTTTGGCTCTCTGAGCACTTGGCTCT 1728
DB 1526 NNNNGACAGCTTGTGAGGCTCGGGCCGCCACACACTTGGCTCTTGTGAAGAACTGGCC 1585
QY 1729 CTTCCATCTCTGAGAGCCCGGCGCAGCGCCCTTCAAGGGCTGTGAGTGAAGACGATCC 1788

Db 1586 CATCTGCCCCCTGCGCCGCGGCGAGCCGACCTCAAGGGCTGTGAGCGAGGACAGCATCC 1645
Qy 1789 TGTCTCCGAGTCTTGTGACCAATTTGACCTTCCGTAACGTCCTTCCGAAACCCCACTGA 1848
Db 1646 TGTCTCTGAGTCTTGTGACCAAGCTTGAAGGCTTCCGTAACGTCCTTCCGAAACCCCACTGC 1705
Qy 1849 GGGGCTGTGTCTGTGTGACCAAGCTTGAAGGCTTGAAGGCTTCCGTAACGTCCTTCCGTAACG 1902
Db 1706 GGGGCTGTGTCTGTGTGACCAAGCTTGAAGGCTTGAAGGCTTCCGTAACGTCCTTCCGTAACG 1765
Qy 1903 -----GTCTGAACGATGTGTGAGCAAGATCTTGGGGGATAGCTGCTTTTCTGTGACG 1956
Db 1766 GAACTGCTGTGAGCGCTGTGCGGAGGATCTTGTGGGAGCAAGCTGCTTTTCTGTGACG 1825
Qy 1957 ACTGCGAAGGATGATCTGACCTTCAAGCAAGGCTTGAAGATGCTCAAGGCTCAAGCT 2016
Db 1826 ACTGCGAAGGATGATGACGCTTCAAGCAAGGCTTGAAGATGCTCAAGGCTCAAGCT 1885
Qy 2017 GA 2018
Db 1886 GA 1887

RESULT 8
CR749209 7004 bp mRNA linear HTC 19-AUG-2004
LOCUS Homo sapiens mRNA; cDNA DKFP686F0113 (from clone DKFP686F0113).
DEFINITION CR749209
ACCESSION CR749209.1 GI:51476151
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 7004)
Ostenmaeder,B., Obermaier,B., Deutscherhaur,S., Schaijpp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M., and Wiemann,S.
The German cDNA Consortium
CONSTRM Direct Submission
TITLE Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
JOURNAL Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by Medigenomix (Martinsried/Germany) within the cDNA Sequencing consortium of the German Genome Project.
COMMENT This clone (DKFP686F0113) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFP686F0113
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
source Location/Qualifiers
1..7004
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686F0113"
/tissue_type="fetal kidney"
/clone_lib="686 (synonym: hlc3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"
/dev_stages="fetal"
/note="hypothetical protein, N-terminus truncated, not fully spliced"
1..7004
/gene="DKFP686F0113"
<4163..5479
/gene="DKFP686F0113"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18066.1"
/db_xref="GI:51476152"
/translation="TADPGLSNLYHGKFLQTFGCSPLVASEIVNGKPYTPYDVS"

SLGLVLYILVHGTMPDGHDKILVQIISNGVAREPKPSDAGLIMLVNPTERRA
TLBDAVSHMMWMMGYATRVGEQAEHFGHGPDSASAMADWLRSSRLPLENGARV
CSFEKQAPGGSTTPGLERQSLKSKRKNEMQSHSDPADTARPGKSNLKP
GILKKKYSABGVQEDPEPLSPAPAGGAPLLPKGILKKROESGYSPEPS
ESSELDAAGGVQSDPEKRPDASDILLHRGKILKLNKESQTALEIAPTFGSL
DELAPEPLARASPGASVEDSILSSESDQDLPLRLPEPLRGCVSDNLTGLES
PSEBGPSCILRWRMDPLDPSCSFLDQCEVATYRQALRVCSKLTP

ORIGIN
Query Match 32.2%; Score 934.4; DB 3; Length 7004;
Best Local Similarity 80.6%; Pred. No. 5.7e-209;
Matches 1137; Conservative 0; Mismatches 251; Indels 22; Gaps 3;
695 AATCATCAAGATTCTGATCTTGTGCTCTCCAACTGTACCAAGGCAAGTCTTCA 754
Db 4153 AATCTGCAATTCCTGATCTTGTGCTCTCCAACTGTACCAAGGCAAGTCTTCA 4212
Qy 755 GACGTTCTGTGAGGACCTCTGTACGCTGCTGTGAGATGATCAAGGAAAGCTTATGT 814
Db 4213 GACATTCGTGTGAGGACCTCTGTACGCTGCTGTGAGATGATGATGAGGAAAGCTTACAC 4272
Qy 815 GGGCCCAAGATGTGACAGCTGTGTCTGTGGGCTTCTCTGTACATCTCTGTGTGATGAC 874
Db 4273 AGGCCAAGATGTGACAGCTGTGTCTGTGGGCTTCTCTGTACATCTCTGTGTGATGAC 4332
Qy 875 CATGCCCTTTGACGGGAGGATCATTAACACTGATGAGCAATCATGTAAGGGGCTTCA 934
Db 4333 CATGCCCTTTGATGAGGATCATTAACACTGATGAGCAATCATGTAAGGGGCTTCA 4392
Qy 935 CCGTGAACCGGCCCAAGCGCTGTGATGCTGTGAGCTGTATCCGGTGTGATGATGATGAT 994
Db 4393 CCGGAGAGCACTTAACCGCTGTGATGCTGTGAGCTGTATCCGGTGTGATGATGATGAT 4452
Qy 995 CCCCACCCGTCGGGCCCACTGAGGATGATGAGCAATGATGATGATGATGATGATGATGAT 1054
Db 4453 CCCCACCCGTCGGGCCCACTGAGGATGATGAGCAATGATGATGATGATGATGATGATGAT 4512
Qy 1055 CACCACCGGATGAGGAGGAGCAAGAGCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1114
Db 4513 CCGCACCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4572
Qy 1115 TGCCGCGGCTCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1174
Db 4573 TGCCGCGGCTCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4632
Qy 1175 AGCCAAAGTGTGACGCTTCTTCAAGCAAGCAAGTGTGAGGAGGAGGAGGAGGAGGAGG 1234
Db 4633 GGCACAAAGTGTGACGCTTCTTCAAGCAAGCAAGTGTGAGGAGGAGGAGGAGGAGGAGG 4692
Qy 1235 GCTGAGCGGCAATCTCTTAAAGAGTCCGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1294
Db 4693 CTTGAGCGGCAATCTCTTAAAGAGTCCGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4752
Qy 1295 GCAAGGTGACCGGCTGAGGATGATCTCTTCTGCGCTGTGCAAGAGGAGGAGGAGGAGG 1354
Db 4753 CCAAGGTGACCGGCTGAGGATGATCTCTTCTGCGCTGTGCAAGAGGAGGAGGAGGAGGAGG 4812
Qy 1355 GAAAGGCAATCTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1414
Db 4813 AAGAGGCAATCTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4872
Qy 1415 GGAATCTGAGCGGCTGTGATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1474
Db 4873 GGAATCTGAGCGGCTGTGATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4932
Qy 1475 AAGGAAAGGCAATCTTAAAGATGATGACAGCGGATGATGATGATGATGATGATGATGAT 1534
Db 4924 CAAGAAGGCAATCTTAAAGATGATGACAGCGGATGATGATGATGATGATGATGATGATGAT 4983
Qy 1535 GCCCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594
Db 4984 GCCCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5043

/dev stage="embryo.12.5.13.5.14.5 dpc"
/lab host="DH10B (T1 phage resistant)"
/clone lib="NTH_BMAP_HDO"
/note="Organ: Eye; Vector: pYX-Abs; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Abs vector. The library tag
sequence located between the Not I site and the polyA tail
is TTAATGAAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
Developing Mouse Nervous System", supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 24.6%; Score 714; DB 7; Length 727;
Best Local Similarity 99.6%; Pred. No. 3.6e-15; Mismatches 2; Indels 1; Gaps 1;
Matches 725; Conservative 0;

Db 1828 GTCTTCCGAAACCCCACTGAGGAGCTGTGTCTGTGAGACACCTGAGGAGCTTGAAC 1887
1 GTCTTCCGAAACCCCACTGAGGAGCTGTGTCTGTGAGAC-ACCTGAGGAGCTTGAAC 59
Qy 1888 AGCTTCCCTCGAAGAGTCTGAAGGATGTGTGCGAGAAATCCTTGGGGGATAGCTGCTTTT 1947
60 AGCTTCCCTCGAAGAGTCTGAAGGATGTGTGCGAGAAATCCTTGGGGGATAGCTGCTTTT 119
Db 1948 CTCTGACGAGCTGCAAGAGGTGACTGAGGCTTACAGACAGCCCTGAGATTCGCTCAA 2007
120 CTCTGACGAGCTGCAAGAGGTGACTGAGGCTTACAGACAGCCCTGAGATTCGCTCAA 179
Qy 2008 AGCTCAAGTGAAGAGGAGATGTGTGCTGATGTGGGTAGGCTCTGAGAGGCTTGA 2067
180 AGCTCAAGTGAAGAGGAGATGTGTGCTGATGTGGGTAGGCTCTGAGAGGCTTGA 239
Db 2068 GAGAAACCTTGGTTCGATTTCTTCAAGTGAATGATGATCAATCAAGGCTTACGCTTGA 2127
240 GAGAAACCTTGGTTCGATTTCTTCAAGTGAATGATGATCAATCAAGGCTTACGCTTGA 299
Qy 2128 GCTTGAAGTGAAGTGAAGGATGAGAAATGCAATGATGAGAAAGGAAAGGAAACCTT 2187
300 GCTTGAAGTGAAGTGAAGGATGAGAAATGCAATGATGAGAAAGGAAAGGAAACCTT 359
Db 2188 TGCTGCGGAGTGTATAGTGGGGTGTGCTGAGAGTGCCTTCTTTGTGCATGAT 2247
360 TGCTGCGGAGTGTATAGTGGGGTGTGCTGAGAGTGCCTTCTTTGTGCATGAT 419
Qy 2248 GTTCAACCATGATTTTCCACCCCTGTTCTGTGCTGACACTTCAATTAATTTCTGTTTC 2307
420 GTTCAACCATGATTTTCCACCCCTGTTCTGTGCTGACACTTCAATTAATTTCTGTTTC 479
Db 2308 CATCAACCAACGAGGTTTGAACCCCTGATTCCTGGGAGTAAATGTATGATGATGATTC 2367
480 CATCAACCAACGAGGTTTGAACCCCTGATTCCTGGGAGTAAATGTATGATGATGATTC 539
Qy 2368 ATTATGAGAGGAAACAGCTCTGTGTTTCATCTGTGCTGTGCTCATCAAGACCTGG 2427
540 ATTATGAGAGGAAACAGCTCTGTGTTTCATCTGTGCTGTGCTCATCAAGACCTGG 599
Db 2428 GAAAGCTCGGACCGCTGTTGATTTCAATCTCAAGGGGACAGATGCCCTGAGACCCATC 2487
600 GAAAGCTCGGACCGCTGTTGATTTCAATCTCAAGGGGACAGATGCCCTGAGACCCATC 659
Qy 2488 TTATGATCTCAGAGCTTGAACCTTGAAGCTTTCTTATGATCCGAGATGTGATGATGCT 2547
660 TTATGATCTCAGAGCTTGAACCTTGAAGCTTTCTTATGATCCGAGATGTGATGATGCT 719
Db 2548 CTGTTTCT 2555
Qy

Db 720 CTGTTTCT 727

RESULT 12
BG915967
LOCUS
DEFINITION
60281523.F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4937647 5',
mRNA sequence.
ACCESSION
BG915967
VERSION
BG915967.1 GI:14296443
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 992)
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strauberg, Ph.D.
Email: Ggabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM10871 row: m column: 08
High quality sequence stop: 695.
Location/Qualifiers
1. 992
/organism="Mus musculus"
/mol type="mRNA"
/strain="NMRI"
/db xref="taxon:10090"
/clone="IMAGE:4937647"
/tissue="tumor, gross tissue"
/dev stage="5 months"
/lab host="DH10B"
/clone lib="NCI CGAP Mam4"
/note="Organ: mammary; pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furch,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

FEATURES

source

ORIGIN

Query Match 24.1%; Score 700.2; DB 4; Length 992;
Best Local Similarity 92.2%; Pred. No. 6.9e-154; Mismatches 857; Conservative 0; Indels 12; Gaps 11;

Qy 1931 GGGGATAGTGTCTTCTCTGACAGCTGCAAGAGTGTGAGGCTTCAAGACAGC 1990
13 GGGGATAGTGTCTTCTCTGACAGCTGCAAGAGTGTGAGGCTTCAAGACAGC 72
Db 1991 CCTAGGAATCTGCTCAAGAGTCAAGTGAAGGAGATGTGAGGCTTCAAGAGG 2050
73 CCTAGGAATCTGCTCAAGAGTCAAGTGAAGGAGATGTGAGGCTTCAAGAGG 132
Qy 2051 CTCTGAGAGGTTTGCAGAGAAACCTTGGTTCGATTTCTTCAAGTATGAGTATCA 2110
133 CTCTGAGAGGTTTGCAGAGAAACCTTGGTTCGATTTCTTCAAGTATGAGTATCA 192
Db 2111 AGGGCTTACGTCTGACAGCTGAACTGAAAGTGAAGAAATGCAATGATGATG 2170
193 AGGGCTTACGTCTGACAGCTGAACTGAAAGTGAAGAAATGCAATGATGATG 252
Qy 2171 AAAGGAATGGAAACCTTGTGCGCCGAGTGTATGAGGAGTGGCCGAGAGTGCAC 2230
253 AAAGGAATGGAAACCTTGTGCGCCGAGTGTATGAGGAGTGGCCGAGAGTGCAC 312
Db

Db 613 GCCACGAGAGGAGATGTTTACCGAGATCTTCAGCTGGAGAAACATCTTCTAATGCCA 672

QY 691 ATGGAACATCAAGATTGCTGACTTTGGCCTCT 723

Db 673 NTGGAAACATCANGATTGCTGACTTTGGACTCT 705

RESULT 14

CB248251

LOCUS

DEFINITION CB248251 726 bp mRNA linear EST 15-JUL-2003

IMAGE:5718428 5', mRNA sequence.

CB248251

VERSION CB248251.1 GI:28386432

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 726)

NH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAN at: <http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

Location/Qualifiers

1. 726

/organism="Mus musculus"

/mol_type="mRNA"

/screen="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5718428"

/issue_type="whole brain"

/dev_stage="embryo 12.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_id="NH BMAP FDO"

/note="Organ: brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lemmon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.6%; Score 683.6; DB 6; Length 726;

Best Local Similarity 98.8%; Pred. No. 5.3e-150;

Matches 720; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 922 GTAAAGGGGCTTAC-GGTGAGCCGCCCAAGCGTCCGATGCTGTGGCTGATCCGGTGG 980

Db 1 GTAAAGGGGCTTACGCTGAGCCGCCCAAGCGTCCGATGCTGTGGCTGATCCGGTGG 60

QY 981 CTGTATATGTGAACCCACCCGCTGGGCCACTGAGAGATGTAGCCAGTCATTGGTGG 1040

Db 61 CTGTATATGTGAACCCACCCGCTGGGCCACTGAGAGATGTAGCCAGTCATTGGTGG 120

QY 1041 GTCAACTGGGTTACACACCGAGTCCGGGAAACAGAAACCCCTGCGTGAAGGTGGAC 1100

Db 121 GTCAACTGGGTTACACACCGAGTCCGGGAAACAGAAACCCCTGCGTGAAGGTGGAC 180

QY 1101 CTTAGTGTGATCTTTGGCCGGGCTCCATGCGGAGCTGGTTAGTGCCTCTCGCGCCC 1160

Db 181 CTTAGTGTGATCTTTGGCCGGGCTCCATGCGGAGCTGGTTAGTGCCTCTCGCGCCC 240

QY 1161 CTCTGGAAGATGAGACCAAGTGTGAGCTTCTTCAAGAGACGCGGGAGGTGA 1220

Db 241 CTCTGGAAGATGAGACCAAGTGTGAGCTTCTTCAAGAGACGCGGGAGGTGA 300

QY 1221 AGCACTGTACCTGGGCTGAGCGGCAACATTTCTTAAAGTCCCGAAAGAGATGAC 1280

Db 301 AGCACTGTACCTGGGCTGAGCGGCAACATTTCTTAAAGTCCCGAAAGAGATGAC 360

QY 1281 ATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGAGATCTCTTCTGCGCTGGCAAGAC 1340

Db 361 ATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGAGATCTCTTCTGCGCTGGCAAGAC 420

QY 1341 AGCTTAAAGCTTCGAAAGGCAATTCAGAAAGGCTCTACCTCTGAGGGGAGTA 1400

Db 421 AGCTTAAAGCTTCGAAAGGCAATTCAGAAAGGCTCTACCTCTGAGGGGAGTA 480

QY 1401 CAGAGAGACCTTCAGAGAACTCAGACCGGTGCTGATTAATCTCAGAGGACCTGTCCCTGT 1460

Db 481 CAGAGAGACCTTCAGAGAACTCAGACCGGTGCTGATTAATCTCAGAGGACCTGTCCCTGT 540

QY 1461 GTATCTCTGCTCCCAAGAAAGGCAATCTTAAAGTCTGACAGCTGTAATCTGTTAC 1520

Db 541 GTATCTCTGCTCCCAAGAAAGGCAATCTTAAAGTCTGACAGCTGTAATCTGTTAC 600

QY 1521 TACTCTCTTCAGAGGCTCAGAGGCTGCGGGAACTCTTAAAGGCTGATGTTGTG 1580

Db 601 TACTCTCTTCAGAGGCTCAGAGGCTGCGGGAACTCTTAAAGGCTGATGTTGTG 660

QY 1581 AGTGGGAGCCCGGTGAGCAGAAAGTCTCAGAGGCTTCTGCTTCCACCGCAAG 1640

Db 661 AGGAGGAA-CCCGTGGAGCAGAA--TCTCAGAGGCTTCAAGGGCTCTCTCCACCGCAAG 717

QY 1641 GGCATTCTC 1649

Db 718 GGCATTCTC 726

RESULT 15

CO424322

LOCUS

DEFINITION CO424322 672 bp mRNA linear EST 06-JUL-2004

IMAGE:30665189 5', mRNA sequence.

CO424322

VERSION CO424322.1 GI:49670481

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 672)

NH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at <http://genome.iowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. .672
source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:3065389"
/tissue_type="whole eye"
/dev_stage="newborn (1, 5, 15 days)"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="NIH_BMAP_H00"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AATAATGCG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 23.2%; Score 672; DB 7; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.8e-147;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 913 AGCAATCAGTAAGGGGCTTACCGTAGCGGCCCAAGCGTCCGATGCTGTGGCCCTGA 972
Db 1 AGCAATCAGTAAGGGGCTTACCGTAGCGGCCCAAGCGTCCGATGCTGTGGCCCTGA 60
QY 973 TCCGGTGGCTGTATGTGTGAACCCCAACCCCTCGGGCCACACTGAGAGATGTAGCAATC 1032
Db 61 TCCGGTGGCTGTATGTGTGAACCCCAACCCCTCGGGCCACACTGAGAGATGTAGCAATC 120
QY 1033 ATTGGTGGGTCACTGGGGTTACCAACCGGAGTGGGGAAACAGAAAGCCCTGCGTAGG 1092
Db 121 ATTGGTGGGTCACTGGGGTTACCAACCGGAGTGGGGAAACAGAAAGCCCTGCGTAGG 180
QY 1093 GTGGGCAACCCCTAGGTGTGACTTTGGCCGGGCTCCATGCGGACTGTTACGTGCTCCT 1152
Db 181 GTGGGCAACCCCTAGGTGTGACTTTGGCCGGGCTCCATGCGGACTGTTACGTGCTCCT 240
QY 1153 CGGCCCCCTCTGGAAGATGAGCAAGTGTGCAAGCTTTCAGAGCAGACGTGCCGG 1212
Db 241 CGGCCCCCTCTGGAAGATGAGCAAGTGTGCAAGCTTTCAGAGCAGACGTGCCGG 300
QY 1213 GAGGTGGAAGCACTGTACTGTGGGCTGAGCGGCAACATTTCTTAAGAAATCCGAAAGG 1272
Db 301 GAGGTGGAAGCACTGTACTGTGGGCTGAGCGGCAACATTTCTTAAGAAATCCGAAAGG 360
QY 1273 AGAATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTTTCTGCGCCTG 1332
Db 361 AGAATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTTTCTGCGCCTG 420
QY 1333 GCAAGAGAGCCTTAAGCTTCGAAAGGCAATTCAGAAAGAAAGCTCTACCTCGTAG 1392
Db 421 GCAAGAGAGCCTTAAGCTTCGAAAGGCAATTCAGAAAGAAAGCTCTACCTCGTAG 480
QY 1393 GGAAGGTAACAGAGAGACCTTCAGAAATCTCAGACCGGTGCTGATACCTCAGGGAGCCTG 1452
Db 481 GGAAGGTAACAGAGAGACCTTCAGAAATCTCAGACCGGTGCTGATACCTCAGGGAGCCTG 540
QY 1453 TCCCTGCTGTATCCCTGCTCCAGAGAAAGGATCCTTAAGAGTCTGACAGCGTAGAT 1512
Db 541 TCCCTGCTGTATCCCTGCTCCAGAGAAAGGATCCTTAAGAGTCTGACAGCGTAGAT 600
```

```
QY 1513 CTGATTACTACTCTCTTCCAGAGCCAGCCAGGATCTGGGGAACCTTTAGACCCAGTAGTG 1572
Db 601 CTGATTACTACTCTCTTCCAGAGCCAGCCAGGATCTGGGGAACCTTTAGACCCAGTAGTG 660
QY 1573 TGTTTGAGTGG 1584
Db 661 TGTTTGAGTGG 672
```

Search completed: May 16, 2005, 03:15:24
Job time : 9317.58 secs

THIS PAGE BLANK (USPTO)

Db 61 CCCGCGCTCAGCGCGCTGCTACTGCTGCCGACCACTCACTTCGCGGTCCCCGCA 120
QY 121 CCATGAGTGGTGGCTTACTCCAGGCGCCGAGACGAGCTCCTCGGCTCCGCTCGG 180
Db 121 CCATGAGTGGTGGCTTACTCCAGGCGCCGAGACGAGCTCCTCGGCTCCGCTCGG 180
QY 181 CCTCGAGAGCGCCGCGCTGCGGACGCGGCTCATCAAGTCGCTTAACTCTGATGA 240
Db 181 CCTCGAGAGCGCCGCGCTGCGGACGCGGCTCATCAAGTCGCTTAACTCTGATGA 240
QY 241 AGAAGCAGCGGTGAAGCGGACCATCAAAACAACCTGCGGCACTGCTACGAGTTCC 300
Db 241 AGAAGCAGCGGTGAAGCGGACCATCAAAACAACCTGCGGCACTGCTACGAGTTCC 300
QY 301 TGAAGAGCTGGGGAAGAGGCACTTACGAGAAAGGTGAAGAGGCAAGAGAGCTCGGCGC 360
Db 301 TGAAGAGCTGGGGAAGAGGCACTTACGAGAAAGGTGAAGAGGCAAGAGAGCTCGGCGC 360
QY 361 GTCGTGAGCCATCAATCCATCAGAAAGAACAATCAAGATGAGAGGATCGCTGC 420
Db 361 GTCGTGAGCCATCAATCCATCAGAAAGAACAATCAAGATGAGAGGATCGCTGC 420
QY 421 ACATACGAGGAGAGATTGAATCATGTCTTCACTCAACCAACCCCACTATCCATCC 480
Db 421 ACATACGAGGAGAGATTGAATCATGTCTTCACTCAACCAACCCCACTATCCATCC 480
QY 481 ATGAAGTTTGAAGAAATGACGCAAGATGTGATTCATGAGATGACGAGCGG 540
Db 481 ATGAAGTTTGAAGAAATGACGCAAGATGTGATTCATGAGATGACGAGCGG 540
QY 541 ATCTGATGATTAATCATGATGAGCGGCAAGCTGATGAGCGGCAAGCGGATTTCT 600
Db 541 ATCTGATGATTAATCATGATGAGCGGCAAGCTGATGAGCGGCAAGCGGATTTCT 600
QY 601 TCCGACGATGCTGTCTGCGCTGACATGCGCAACGAGGATGTTCAACCGAATC 660
Db 601 TCCGACGATGCTGTCTGCGCTGACATGCGCAACGAGGATGTTCAACCGAATC 660
QY 661 TCAAGCTGGAAGAAATCTCTTCTGATGTCGCAATGAGAAATCAAGATGCTGATG 720
Db 661 TCAAGCTGGAAGAAATCTCTTCTGATGTCGCAATGAGAAATCAAGATGCTGATG 720
QY 721 TCTCCAACTGTACCAAAAGGCAAGTCTCCAGAGCTTCTGAGGAGCTCTACG 780
Db 721 TCTCCAACTGTACCAAAAGGCAAGTCTCTCCAGAGCTTCTGAGGAGCTCTACG 780
QY 781 CCTGCGCTGAGATGATCAACGAGGAGCTTATGTGAGCCAGAGGTGAGACGCTGCTC 840
Db 781 CCTGCGCTGAGATGATCAACGAGGAGCTTATGTGAGCCAGAGGTGAGACGCTGCTC 840
QY 841 TGGGCGCTTCTCTGTATCATCTGATGATGAGCAACATGCTTGAAGGAGATCAT 900
Db 841 TGGGCGCTTCTCTGTATCATCTGATGATGAGCAACATGCTTGAAGGAGATCAT 900
QY 901 AAACACTGTGAAGCAAAATCATGTAACGAGGCTTACGAGCGGCAAGCGCTGATG 960
Db 901 AAACACTGTGAAGCAAAATCATGTAACGAGGCTTACGAGCGGCAAGCGCTGATG 960
QY 961 CCTGTGAGCTGATCCGCTGAGCTGTTAATGTGAACCCCAACCGCTGAGGAGG 1020
Db 961 CCTGTGAGCTGATCCGCTGAGCTGTTAATGTGAACCCCAACCGCTGAGGAGG 1020
QY 1021 ATGTAGCAATCATTTGTGTGATCACTGAGGTTTACACACCGAGTGGGGAACGAG 1080
Db 1021 ATGTAGCAATCATTTGTGTGATCACTGAGGTTTACACACCGAGTGGGGAACGAG 1080
QY 1081 CCCGCGTGAAGGAGGAGCCATGATGATGCTTGGCGGCGCTCATGCGGAGCTGCT 1140
Db 1081 CCCGCGTGAAGGAGGAGCCATGATGATGCTTGGCGGCGCTCATGCGGAGCTGCT 1140
QY 1141 TACGTGCTCTCGCGCCCTCTCTGAGAGATGAGCAAGTGTGACGCTTCTTCAAGC 1200
Db 1141 TACGTGCTCTCGCGCCCTCTCTGAGAGATGAGCAAGTGTGACGCTTCTTCAAGC 1200

QY 1201 AGCACGTGCGCGGAGAGTGAAGCATCTGTACTGTGGCTGAGAGCGGCAACATTTCTTAA 1260
Db 1201 AGCACGTGCGCGGAGAGTGAAGCATCTGTACTGTGGCTGAGAGCGGCAACATTTCTTAA 1260
QY 1261 AGTCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 AGTCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 CTCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCT 1380
Db 1321 CTCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCT 1380
QY 1381 CTACTCTGTCAAGGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 CTACTCTGTCAAGGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 CAGGAG 1500
Db 1441 CAGGAG 1500
QY 1501 GACAGCGTGAATCTGTATCTACTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 GACAGCGTGAATCTGTATCTACTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 AGCGCAATGATGTGTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 AGCGCAATGATGTGTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 GGCCTCTCTCAACCGGAG 1680
Db 1621 GGCCTCTCTCAACCGGAG 1680
QY 1681 TAGAAGCACTAACCCCTAGACCTTGTGCTCCCTGAGCAACTGAGCTCTCCATCTG 1740
Db 1681 TAGAAGCACTAACCCCTAGACCTTGTGCTCCCTGAGCAACTGAGCTCTCCATCTG 1740
QY 1741 CAGCGCGGAG 1800
Db 1741 CAGCGCGGAG 1800
QY 1801 CTTTGAACAAATTTGAACTTGTGCTGAAAGTCTTCCGAAACCCCACTGAGGAGCTG 1860
Db 1801 CTTTGAACAAATTTGAACTTGTGCTGAAAGTCTTCCGAAACCCCACTGAGGAGCTG 1860
QY 1861 CTGTGACAACTGAGGAG 1920
Db 1861 CTGTGACAACTGAGGAG 1920
QY 1921 AGGAATCTTGGGAGATGCTGCTTCTCTGACAGCTGCAAGAGTGAATGAGAGCT 1980
Db 1921 AGGAATCTTGGGAGATGCTGCTTCTCTGACAGCTGCAAGAGTGAATGAGAGCT 1980
QY 1981 ACAAGCAAGGCTGAGAGATCTGCTCAAGGCTGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 1981 ACAAGCAAGGCTGAGAGATCTGCTCAAGGCTGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 2041 ATGGGATGAGCTGTGAG 2100
Db 2041 ATGGGATGAGCTGTGAG 2100
QY 2101 GAGTACATCAAGGAGCTTACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 GAGTACATCAAGGAGCTTACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2161 ATTGATGTGAAG 2220
Db 2161 ATTGATGTGAAG 2220
QY 2221 GGTGCTACTCTCTTGTGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 GGTGCTACTCTCTTGTGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 2280

QY	2281	TTGACACCTTCACATAAGTTTCGTGTTTCATCAACCAACAGGGGTTAGAACCTGTACTCT	2340
Db	2281	CTGACCTTCACATBAAGTTTCGTGTTTCATCAACCAACAGGGGTTAGAACCTGTACTCT	2340
QY	2341	GGAGAGTATGTAGTACCTGCCATTATTTAGAGAGGAAAAGAGCTCTGTGTTTCATCT	2400
Db	2341	GGAGAGTATGTAGTACCTGCCATTATTTAGAGAGGAAAAGAGCTCTGTGTTTCATCT	2400
QY	2401	CTGCTGCTGTGTCATCTCAAAAGACCTGGGAAAGACTGGACCGCTGTTTGACTTCAATCTCAA	2460
Db	2401	CTGCTGCTGTGTCATCTCAAAAGACCTGGGAAAGACTGGACCGCTGTTTGACTTCAATCTCAA	2460
QY	2461	GGGGACCAAGATGCCCTCGAACCCCATCTTATGTCAGAGACTGGAACCTTGAAGCTGTT	2520
Db	2461	GGGGACCAAGATGCCCTCGAACCCCATCTTATGTCAGAGACTGGAACCTTGAAGCTGTT	2520
QY	2521	CCTAGTACCCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTGC	2580
Db	2521	CCTAGTACCCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTGC	2580
QY	2581	TGACTTATTTATTTTTTTTGTGATTTCTCACTTCTGTGTTTTTGGTTTTTGTGTTTGG	2640
Db	2581	TGACTTATTTATTTTTTTTGTGATTTCTCACTTCTGTGTTTTTGGTTTTTGTGTTTGG	2640
QY	2641	TTTTTGTTTTTAACTGAATTTTGTGCTGCTTTCATAAATAGTGAATCTGTGTTCTGCGGAAC	2700
Db	2641	TTTTTGTTTTTAACTGAATTTTGTGCTGCTTTCATAAATAGTGAATCTGTGTTCTGCGGAAC	2700
QY	2701	TCCACTGTGCACTGAAGTTTATGTATGTAACAGAAAGTATTTGGCAATGATGTCCCTCTATTC	2760
Db	2701	TCCACTGTGCACTGAAGTTTATGTATGTAACAGAAAGTATTTGGCAATGATGTCCCTCTATTC	2760
QY	2761	AAGGGGGGTGGGGGCGTTTTTCAAAATGTATGCTTGGACACTGTCTGGATTGAATCTCCA	2820
Db	2761	AAGGGGGGTGGGGGCGTTTTTCAAAATGTATGCTTGGACACTGTCTGGATTGAATCTCCA	2820
QY	2821	GTCCCTTCACACCCCAAGGCTGGCCACCTCCCTCATCTTCACTGTGTGGCCAAAAAAA	2880
Db	2821	GTCCCTTCACACCCCAAGGCTGGCCACCTCCCTCATCTTCACTGTGTGGCCAAAAAAA	2880
QY	2881	AAAAAAAAAAAAAAAAAAAAA 2902	
Db	2881	AAAAAAAAAAAAAAAAAAAAA 2902	

Query	Match	Similarity	98.7%	Score	286.4	DB	18	Length	3073
Best Local	Similarity	100.0%	Pred.	No.	0				
Matches	2864	Conservative	0	Mismatches	1	Indels	0	Gaps	0

Qy	32	CGTGTCTGGGATGCGGCTGTGACCTTCTGAGCCCCGCGCTCAACGGCGGCTGTACTTACTGCTGC	91
Db	1	CTGTGTCTGGGATGCGGCTGTGACCTTCTGAGCCCCGCGCTCAACGGCGGCTGTACTTACTGCTGC	60
Qy	92	CGAGCCACTCACCCTCGGGTCCCGGACCATGAGTGGTGGACCTTACTCCAGGCC	151

[illegible]

Db 1141 TGGAGCCAAGGTGTGAGCTTCTTCAAGCAGACGTGCGGAGGTGGAACACTGTACC 1200
Qy 1232 TGGGCTGAGCGGGAACATTTCTTTAAGAGTCCCGAAGAGATGACATGGCTCAAA 1231
Db 1201 TGGGCTGAGCGGGAACATTTCTTTAAGAGTCCCGAAGAGATGACATGGCTCAAA 1260
Qy 1292 TCTGCAAGTGAACCGGCTGAGATACCTTTCTCGCCCTGGCAAGAGACCTTTAAGT 1351
Db 1261 TCTGCAAGTGAACCGGCTGAGATACCTTTCTCGCCCTGGCAAGAGACCTTTAAGT 1320
Qy 1352 TCGAAGAGCATTTCTCAAGAAAAGTCTTACCTCGTCAAGGAGGTACAGAGACC 1411
Db 1321 TCGAAGAGCATTTCTCAAGAAAAGTCTTACCTCGTCAAGGAGGTACAGAGACC 1380
Qy 1412 TCGAAGCATGACCGGCTGATTAATCTCCAGGGGACGCTTCCCTGTGTATCCCTGCT 1471
Db 1381 TCGAAGCATGACCGGCTGATTAATCTCCAGGGGACGCTTCCCTGTGTATCCCTGCT 1440
Qy 1472 CCCAAGGAAGGATCTTAAAGTCTGACAGCGTGAATCTGGTTACTACTCTCTCC 1531
Db 1441 CCCAAGGAAGGATCTTAAAGTCTGACAGCGTGAATCTGGTTACTACTCTCTCC 1500
Qy 1532 AGAGCCGAGGCTGAGGAACTTTAGACGCGAGTGTGTGTGAGTGGGAGCC 1591
Db 1501 AGAGCCGAGGCTGAGGAACTTTAGACGCGAGTGTGTGTGAGTGGGAGCC 1560
Qy 1592 CGTGAACAGAGTCTCAAGAGGTTCAAGGCTCTCTCTCAAGGAGGATTTTCA 1651
Db 1561 CGTGAACAGAGTCTCAAGAGGTTCAAGGCTCTCTCTCAAGGAGGATTTTCA 1620
Qy 1652 ACTCAATGGCAAGTCTCCGCAAGGCTTGAAGAGCACTACCCCTAGACCTTTGAGTCT 1711
Db 1621 ACTCAATGGCAAGTCTCCGCAAGGCTTGAAGAGCACTACCCCTAGACCTTTGAGTCT 1680
Qy 1712 CTTGAGCAACTGAGCTCTCTCCATCTGTGACGCGGAGCCGAGCCCTTCAAGGAGTCT 1771
Db 1681 CTTGAGCAACTGAGCTCTCTCCATCTGTGACGCGGAGCCGAGCCCTTCAAGGAGTCT 1740
Qy 1772 GAGTGAAGAGAGTCTCTGCTCCGAGCTCTTGAACCAATTTGAGCTTGAAGTCT 1831
Db 1741 GAGTGAAGAGAGTCTCTGCTCCGAGCTCTTGAACCAATTTGAGCTTGAAGTCT 1800
Qy 1832 TCCGGAACCCCACTGAGGAGGCTGTGTGTGAGCAACCTGAGGGGAGCTTGAAGACC 1891
Db 1801 TCCGGAACCCCACTGAGGAGGCTGTGTGTGAGCAACCTGAGGGGAGCTTGAAGACC 1860
Qy 1892 TCCCTCAAGAGGCTTGAAGGAGTGTGAGCAAGATCTTGGGGAGTACTCTTTCTCT 1951
Db 1861 TCCCTCAAGAGGCTTGAAGGAGTGTGAGCAAGATCTTGGGGAGTACTCTTTCTCT 1920
Qy 1952 GAGAGACTGCGAAGAGTGAAGGCTTGAAGGAGTGTGAGCAAGGCTTGAAGTCT 2011
Db 1921 GAGAGACTGCGAAGAGTGAAGGCTTGAAGGAGTGTGAGCAAGGCTTGAAGTCT 1980
Qy 2012 CAGCTGAGGAAGAGTGTGAGGCTTGAAGGAGTGTGAGGAGGTTTGAAGG 2071
Db 1981 CAGCTGAGGAAGAGTGTGAGGCTTGAAGGAGTGTGAGGAGGTTTGAAGG 2040
Qy 2072 AACCTGGAGTGTGAGTCTCAAGTGAATGATCAATCAAGGAGCTTCAAGTCTGAGACT 2131
Db 2041 AACCTGGAGTGTGAGTCTCAAGTGAATGATCAATCAAGGAGCTTCAAGTCTGAGACT 2100
Qy 2132 GACTGAACCTGAAAGATGAGAAATGCAATGATGAGAAAGAGATGGAGACCTTGTCT 2191
Db 2101 GACTGAACCTGAAAGATGAGAAATGCAATGATGAGAAAGAGATGGAGACCTTGTCT 2160
Qy 2192 GCCGAGTGTATAGTGGGAGGCTGAAGGAGTCTACTCTTGTGAGATAGTGTCA 2251
Db 2161 GCCGAGTGTATAGTGGGAGGCTGAAGGAGTCTACTCTTGTGAGATAGTGTCA 2220
Qy 2252 CCCATGACATTTTCCACCTGTCTCTGAGTCACTTCAATTAAGTTTCTGTTCCATC 2311
Db 2221 CCCATGACATTTTCCACCTGTCTCTGAGTCACTTCAATTAAGTTTCTGTTCCATC 2280

Qy 2312 AACCAAGAGGTTAAGAACCTTGAATCTCTGAGAGTATGTGTAGTACCTGCAATATT 2371
Db 2281 AACCAAGAGGTTAAGAACCTTGAATCTCTGAGAGTATGTGTAGTACCTGCAATATT 2340
Qy 2372 AGAGAGGAACAGGCTGTGGTTTCCATCTGCTGCTGTCATCTCAAGACCTGGGAAG 2431
Db 2341 AGAGAGGAACAGGCTGTGGTTTCCATCTGCTGCTGTCATCTCAAGACCTGGGAAG 2400
Qy 2432 ACTGGAACGCTGTGTTGACTTCAATCTCAAGGAGCAAGATGCCCCCTGAGACCTTAA 2491
Db 2401 ACTGGAACGCTGTGTTGACTTCAATCTCAAGGAGCAAGATGCCCCCTGAGACCTTAA 2460
Qy 2492 ATCTCAAGACTTGAACCTTGAAGCTTCTCTAGTACCAGATGTGATGTGATGCTTGT 2551
Db 2461 ATCTCAAGACTTGAACCTTGAAGCTTCTCTAGTACCAGATGTGATGTGATGCTTGT 2520
Qy 2552 TTCTCAGGCCAAGGAGCCTAGATGTGAGCTTATTTTGTGATTTGATTTCTCACTT 2611
Db 2521 TTCTCAGGCCAAGGAGCCTAGATGTGAGCTTATTTTGTGATTTGATTTCTCACTT 2580
Qy 2612 CTGTTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTT 2671
Db 2581 CTGTTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTTGTGTTTT 2640
Qy 2672 AATTAATGTAATGT 2731
Db 2641 AATTAATGTAATGT 2700
Qy 2732 AAGTATTTGGCAATGATGTCTCTATTTCAAGGAGGAGTGGGAGGTTTTCAATGATG 2791
Db 2701 AAGTATTTGGCAATGATGTCTCTATTTCAAGGAGGAGTGGGAGGTTTTCAATGATG 2760
Qy 2792 TCTTGAAGCACTGTGTGATTAAGTCTTCAAGGAGGAGTGGGAGGTTTTCAATGATG 2851
Db 2761 TCTTGAAGCACTGTGTGATTAAGTCTTCAAGGAGGAGTGGGAGGTTTTCAATGATG 2820
Qy 2852 CTGATCTTCACTGTGTGAGGCAAAAAAATTTTGTGTTTTGTGTTTTGTGTTTTGT 2896
Db 2821 CTGATCTTCACTGTGTGAGGCAAAAAAATTTTGTGTTTTGTGTTTTGTGTTTTGT 2865

RESULT 3
US-10-343-514-1
; Sequence 1, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl P.
; APPLICANT: LEBEYRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: RAT
US-10-343-514-1

Query Match 71.2%; Score 2066.8; DB 18; Length 2929;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2486; Conservative 0; Mismatches 277; Indels 115; Gaps 14;

QY 49 GTGACCTCTGAGCCGCGGCTCAGCGCGCGCTGTACTGTGCCGACCCACTCCACCTC 108
 Db 9 GTGACCTCTGAGCTGCGGCTCTCGCGCGCGCTGTACTGTGCCGACCCACTCCACCTC 68
 QY 109 GCGGTCCTCCCGGACCATGTGAGTGGCTTACTCTCAAGCGCCGAGCCAGGCTCCCTCGG 168
 Db 69 GCGGTCCTCCCGGACCATGTGAGTGGCTTACTCTCAAGCGCCGAGCCAGGCTCCCTCGG 128
 QY 169 CCTCGCGCGCTGCGGACGCGCGCGCGCTGTGCGGAGCGGGCTCATCAAGTCGCGCTA 228
 Db 129 CCTCGCGCGCTGCGGACGCGCGCGCGCTGTGCGGAGCGGGCTCATCAAGTCGCGCTA 188
 QY 229 AACCTCTGATGAAAGAGAGCGCGGTGAGCGGCACTATCAAAACAACACTCTGCGGACC 288
 Db 189 AACCTCTGATGAAAGAGAGCGCGGTGAGCGGCACTATCAAAACAACACTCTGAGGACCC 248
 QY 289 GCTACGAGTTCCTGAGAGCGCTGGGCAAGGGCACTTACGGGAGGTGAGAGAGCGACGAG 348
 Db 249 GCTACGAGTTCCTGAGAGCGCTGGGCAAGGGCACTTACGGGAGGTGAGAGAGCGACGAG 308
 QY 349 AGAGCTGCGGCGCTGTGGTGGCATCAAGTCATCAAGGAGAGAAATCAAGATGAGC 408
 Db 309 AGAGCTGCGGAGCGCTGTGGTGGCATCAAGTCATCAAGGAGAGAAATCAAGATGAGC 368
 QY 409 AGGATCTGTGCACTACGAGGAGATGAGATCATGTCTTCACTCAACCAACCCGACCA 468
 Db 369 AGGATCTGTGCACTACGAGGAGATGAGATCATGTCTTCACTCAACCAACCCGACCA 428
 QY 469 TCATTCGCATTCATGAAAGTGTGTAAGATAGACAGCAAGATTTGATGTCATGAGATAG 528
 Db 429 TCATTCGCATTCATGAAAGTGTGTAAGATAGACAGCAAGATTTGATGTCATGAGATAG 488
 QY 529 CCGAGCGAGCGATCTGTATATATCATCATGAGCGGCGCACTGCTGATGATGAGGCGAGC 588
 Db 489 CCGAGCGAGCGATCTGTATATATCATCATGAGCGGCGCACTGCTGATGATGAGGCGAGC 548
 QY 589 CCGAGCGATTTCTTCCGACAGATCGTGTGCTGCTGCACTATCTGCAACAGAGCGAGTCG 648
 Db 549 CCGAGCGATTTCTTCCGACAGATCGTGTGCTGCTGCACTATCTGCAACAGAGCGAGTCG 608
 QY 649 TTCAACGAGATCTCAAGCTGAGAAACATCTTCTGATGAGTCCCATGAAACATCAAGATG 708
 Db 609 TTCAACGAGATCTCAAGCTGAGAAACATCTTCTGATGAGTCCCATGAAACATCAAGATG 668
 QY 709 CTGACTTTGGGCTCTCCAACTGTATCAAAAGGCAAGTCTCTCAAGACGTTCTGTGGGA 768
 Db 669 CTGACTTTGGGCTCTCCAACTGTATCAAAAGGCAAGTCTCTCAAGACGTTCTGTGGGA 728
 QY 769 GCGCTCTCTACGCGCTGCTGAGATGATCAACGAGGAGCCCTATGTGGGCGCCAGAGGTGG 828
 Db 729 GCGCTCTCTACGCGCTGCTGAGATGATCAACGAGGAGCCCTATGTGGGCGCCAGAGGTGG 788
 QY 829 ACAAGCTGTCTCTGAGGCGTCTCTCTGTATCATCTGTGTGCATGTGACCATGCTTTTGA 888
 Db 789 ACAAGCTGTCTCTGAGGCGTCTCTCTGTATCATCTGTGTGCATGTGACCATGCTTTTGA 848
 QY 889 GGCAGGATCATAAACACTGTGTAAAGCAATCATGTAACGGGGCTTTACCTGTAGCCGCCCA 948
 Db 849 GGCAGGATCATAAACACTGTGTAAAGCAATCATGTAACGGGGCTTTACCTGTAGCCGCCCA 908
 QY 949 AGCGCTCCGATGCTGTGGCGCTGATCCGGTGGCTGTTAAATGTGTAAACCCGCGGTGGG 1008
 Db 909 AGCGCTCCGATGCTGTGGCGCTGATCCGGTGGCTGTTAAATGTGTAAACCCGCGGTGGG 968
 QY 1009 CCACACTGAGAGATGTAGCCAGTCATTTGTGTGTGCTGCACTGGGGTTTACACACCGAGATCG 1068
 Db 969 CCACACTGAGAGATGTAGCCAGTCATTTGTGTGTGCTGCACTGGGGTTTACACACCGAGATCG 1028
 QY 1069 GGGAAACAGAAAGCCCTGTGTAGGGGTGGGCACTTAAAGTGTGATCTTTGGCGCGGCGCTTCA 1128
 Db 1029 GGGAAACAGAAAGCCCTGTGTAGGGGTGGGCACTTAAAGTGTGATCTTTGGCGCGGCGCTTCA 1088

QY 1129 TGGCGAATGTGTTAGTCGCTCTCTCGGCGCCCTCTCTGTGAGATATGAGCCAGGTGTGCA 1188
 Db 1089 TGGCGAATGTGTTAGTCGCTCTCTCTCGGCGCCCTCTCTGTGAGATATGAGCCAGGTGTGTA 1148
 QY 1189 GCTTCTTCAAGCAAGTGTGCGGAGGTGGAAGCATGTATCACTGTGGGTGTGAGGCGGCAAC 1248
 Db 1149 GCTTCTTCAAGCAAGTGTGCGGAGGTGGAAGCATGTATCACTGTGGGTGTGAGGCGGCAAC 1208
 QY 1249 ATTCTCTTAAAGATTCGCCAAGAGATGATCATGTGCTCAAAATCTGCAAGGTGACCCGG 1308
 Db 1209 ATTCTCTTAAAGATTCGCCAAGAGATGATCATGTGCTCAAAATCTGCAAGGTGACCCAG 1268
 QY 1309 CTGAGATACCTCTTCTGCGCCCTGCGCAAGACAGCTTTAAAGTCTCCGAAGGCAATCTCA 1368
 Db 1269 TTGAAGATACCTCTTCTGCGCCCTGCGCAAGACAGCTTTAAAGTCTCCGAAGGTAATCTCA 1328
 QY 1369 AGAAAGATCTCTCACTCTGTCAGGAGGATGATGAGAGAGACCTTCAAGAAATTCAGACCG 1428
 Db 1329 AGAAAGATCTCTCTCTCATGCGGAGGATGATGAGAGAGACCTTCAAGAAATTCAGACCG 1388
 QY 1429 TGCTGATATCTCCAGGCGAGCTGTCTCTGCTGATCCCTGCTCCCAAGAAAGGCAATCC 1488
 Db 1389 TGCTGATATCTCCAGGCGAGCTGTCTCTGCTGATCCCTGCTCCCAAGAAAGGCAATTC 1448
 QY 1489 TTAAGAGTCTCGACAGCGTGAATCTGTGTTACTACTCTCTCTCAAGCGCCAGGATCTG 1548
 Db 1449 TTAAGAGTCTCGACAGCGTGAATCTGTGTTACTACTCTCTCTCAAGCGCCAGGATCTG 1508
 QY 1549 GGGAACTCTTAAGACGCAAGTATGTGTTGTGATGTGGGAGACCCGTGTGAGAGAGATCTC 1608
 Db 1509 GGGAACTCTTAAGACGCAAGTATGTGTTGTGATGTGGGAGACCCGTGTGAGAGAGATCTC 1568
 QY 1609 CACAGGCTTACAGGCGCTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGAGATCT 1668
 Db 1569 CACAGGCTTACAGGCGCTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGAGATCT 1628
 QY 1669 CCGGCAAGCGCTTGAAGAGCACTACCCCTAGACATTTGTGGCTCCCTGAGACCAATGCGCT 1728
 Db 1626 CCGGCAAGCGCTTGAAGAGCACTACCCCTAGACATTTGTGGCTCCCTGAGACCAATGCGCT 1685
 QY 1729 CTTCCATCTCTGACAGCCCGGCGGCGGCGGCTCTGAGGGGCTGTGATGTAGGACAGCATCC 1788
 Db 1686 CTTCCATCTCTGACAGCCCGGCGGCGGCGGCTCTGAGGGGCTGTGATGTAGGACAGCATCC 1745
 QY 1789 TGTCTCGGAGTCTTGTGACCAATTTGACCTTGTGCTGCTGCAAGCGCTTCCGGAACCCCACTGA 1848
 Db 1746 TGTCTCGGAGTCTTGTGACCAATTTGACCTTGTGCTGCTGCAAGCGCTTCCGGAACCCCACTGA 1805
 QY 1849 GGGGCTGTGTCTGTGGAACAACCTGAGGGGGCTTGAAGACGCTTCCCTCAAGAAAGTCTGA 1908
 Db 1806 GGGGCTGTGTCTGTGGAACAACCTGAGGGGGCTTGAAGACGCTTCCCTCAAGAAAGTCTGA 1865
 QY 1909 AGCGATGTGTGAGAGAAATCTTGTGGGGATGAGTGTCTTTCTGTGACAGACTGTCCAGAG 1968
 Db 1866 AGCGATGTGTGAGAGAAATCTTGTGGGGATGAGTGTCTTTCTGTGACAGACTGTCCAGAG 1925
 QY 1969 TGAATGAGAGCTTACAGCAAGCCCTAGAGATCTGCTCAAGCTCAAGCTGAGGAGAGGA 2028
 Db 1926 TGAATGAGAGCTTACAGCAAGCCCTAGAGATCTGCTCAAGCTCAAGCTGAGGAGAGGA 1985
 QY 2029 TGTGCTCTAGT-ATGAGGATGAGCTGTGAGAGGGGTTTGAAGAGAACCTGTGGTCTGAT 2087
 Db 1986 TGTGCTCTAGT-ATGAGGATGAGCTGTGAGAGGGGTTTGAAGAGAACCTGTGGTCTGAT 2044
 QY 2088 CTTCAAGTGAATGATGATCAATCAAG---GAGCTTACGTTGACAGCTGTGATCACTGA 2143
 Db 2045 CTTCAAGTGAATGATGATCAATCAAG---GAGCTTACGTTGACAGCTGTGATCACTGA 2104
 QY 2144 AAGATGAGAGAAATCGATTTGATGTGAGAAAGAAATGAGAAACCTTGTGCGCCGAGGTGTA 2203
 Db 2105 AAGCTGAGAGAAATGAGAGATGTGAGAAAGAAATGAGAAACCTTGTGCGCCGAGGTGTA 2164
 QY 2204 TAGTGGGCT---GAGCTGAAGGTGCTTACCTCTTGTGTGCAATGATGTACCCATGACA 2260

Db 2165 TGAGCGGACACAGAGCTGAAGGTGCTACCTCTTTATGC--TGAGGTGCTACCCATGCCA 2222
 Qy 2261 TTTCCCAACCTGTTCTCTGCTGCAAC-----TTCACTAAAGTTTCTGTTCCATCA 2312
 Db 2223 TCTCCCC-CCCTGCTCTCTGCAAGTGTCAAGGTGTACCACTAATGTTCTGTTGCACTG 2281
 Qy 2313 ACCACAGAGGTGTAACCCCTGACTTCCGAGAGGTAATGTATGATGATGCTGCTATTATTA 2312
 Db 2282 ACCACAGAGGTGTAACCCCTGACTTCCGAGAGGTAATGTATGATGATGCTGCTATTATTA 2341
 Qy 2373 GAGAGAAACAGCTCTGTTGTTTCCATCTCTGCTGCTGCACTCAAGAAGCTGAGAGA 2432
 Db 2342 AAGAGAAACAGCTCTGTTTCCATCTCTGCTGCTGCACTCAAGAAGCTGAGAGA 2401
 Qy 2433 CTGGAACCGCTGTTGATCTTCAAGGAGGACCAAGTCCCTGAGCCCATCTTGA 2492
 Db 2402 CTC-ACTGCTGTTCACTTCATCTCAAGGAGAC----- 2434
 Qy 2493 TCTCAGAGACTTGAACCTTGAAGCTGTTCTAGTACCAGATGATGATGAT- GCTGCT 2551
 Db 2435 -CTCAGAGACTTGAAGCTTGAAGCTGTTCTAGTACCAGATGATGATGATGATGAT 2493
 Qy 2552 TTTCAAGGCCAACCGGAGCTTGAAGTGTGCTGCTTATTTTGTGATTTCTCACTT 2611
 Db 2494 TTCTCAGGCCAGCAGACCCAGATGTGCTGCTTATTTA--TTTGTGATTTCTCACTT 2551
 Qy 2612 CTGTTTTTGGT----- 2624
 Db 2552 CTGTTTCTGTTTCTGTTTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTTT 2611
 Qy 2625 -----TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2618
 Db 2612 TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2611
 Qy 2679 TGAATGCTGTTCTGAGGAACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2738
 Db 2672 TGAATGCTGTTCTGAGGAACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2731
 Qy 2739 TGGCAATGATGTCCTCTATTCAGAGGGGCGTGGGCGTGGGCGTGGGCGTGGGCGT 2798
 Db 2732 TGGCAATGATGTCCTCTATTCAGAGGGGCGTGGGCGTGGGCGTGGGCGTGGGCGT 2783
 Qy 2799 CACTGTGATGATGTCCTCAAGTCCCTTCAACCAAGGCTGGGCGTGGGCGTGGGCGT 2858
 Db 2784 CACTGTGATGATGTCCTCAAGTCCCTTCAACCAAGGCTGGGCGTGGGCGTGGGCGT 2843
 RESULT 4
 US-10-343-514-87
 ; Sequence 87, Application US/10343514
 ; Publication No. US20040132025A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUCKER, Daniel J.
 ; APPLICANT: ROSEN, Cheryl F.
 ; APPLICANT: LEBEYRE, Diana L.
 ; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
 ; FILE REFERENCE: DPA-DRUCZ/PCT
 ; CURRENT APPLICATION NUMBER: US/10/343,514
 ; PRIOR APPLICATION NUMBER: PCT/CA01/01109
 ; PRIOR FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: US 60/222,650
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/274,613
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: CA 2,340,780
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 87
 ; LENGTH: 2027
 ; TYPE: DNA

; ORGANISM: MOUSE
 US-10-343-514-87
 Query Match 69.7%; Score 2022.2; DB 10; Length 2027;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2024; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 49 GTGACCTCTGAGCCCGGCTGAGCGGCGCTGATCTGTCGCCCGGACCCATCTCACTC 108
 Db 1 GTGACCTCTGAGCCCGGCTGAGCGGCGCTGATCTGTCGCCCGGACCCATCTCACTC 60
 Qy 109 GCGGTCCCGCCACCATGAGTCCGTGCTTACTCCAGCCCGGAGCGAGCTCCCTCG 168
 Db 61 GCGGTCCCGCCACCATGAGTCCGTGCTTACTCCAGCCCGGAGCGAGCTCCCTCG 120
 Qy 169 CTTCCGCTGCTGCTGAGAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 228
 Db 121 CTTCCGCTGCTGCTGAGAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 Qy 229 AACCTCTGAG 288
 Db 181 AACCTCTGAG 240
 Qy 289 GCTACGAGTCTTGAAG 348
 Db 241 GCTACGAGTCTTGAAG 300
 Qy 349 AGAGCTCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
 Db 301 AGAGCTCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Qy 409 AGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
 Db 361 AGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Qy 469 TCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
 Db 421 TCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Qy 529 CAGCCGAGGCGATCTGATGATTAATCACTGAGAGCGGCGGAGAGAGAGAGAGAG 588
 Db 481 CAGCCGAGGCGATCTGATGATTAATCACTGAGAGCGGCGGAGAGAGAGAGAGAG 540
 Qy 589 CAGGCACTTTCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
 Db 541 CAGGCACTTTCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Qy 649 TTCAACGAGATCTCAAGCTGAGAAACATCTCTTGAATGCAATGAGAAATCATCAAG 708
 Db 601 TTCAACGAGATCTCAAGCTGAGAAACATCTCTTGAATGCAATGAGAAATCATCAAG 660
 Qy 709 CTGACTTTGCTCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
 Db 661 CTGACTTTGCTCTTCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Qy 769 GCCCTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828
 Db 721 GCCCTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Qy 829 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
 Db 781 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Qy 889 GAGAGATCATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
 Db 841 GAGAGATCATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Qy 949 AGCCGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
 Db 901 AGCCGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Qy 1009 CCAACTGAGAGATGAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068

Db 961 CCAACTGAGAGATGTAGCCAGTATTGTGTGAGTCACTGGGGTTACACCAACCGGATG 1020
Qy 1069 GGGAGACAGAGACCTGTGTGAGGGGTGGGACCTTAGTGTGACTTTGGCGGGCTTCA 1128
Db 1021 GGGAGACAGAGACCTGTGTGAGGGGTGGGACCTTAGTGTGACTTTGGCGGGCTTCA 1080
Qy 1129 TGGGGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1188
Db 1081 TGGGGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
Qy 1189 GCTTCTCAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1248
Db 1141 GCTTCTCAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1200
Qy 1249 ATTCTCTTAAGAGATCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 1308
Db 1201 ATTCTCTTAAGAGATCCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 1260
Qy 1309 CTGAGATACCTCTTCTGCGCTGCGCAAGAGACCTTTAGCTTCCGAAAGGCAATTCGA 1368
Db 1261 CTGAGATACCTCTTCTGCGCTGCGCAAGAGACCTTTAGCTTCCGAAAGGCAATTCGA 1320
Qy 1369 AGAAAGTCTCTACTCTCTCAAGGAGAGTACAGAGAGACCTTCAAGAACTCAAGACCG 1428
Db 1321 AGAAAGTCTCTACTCTCTCAAGGAGAGTACAGAGAGACCTTCAAGAACTCAAGACCG 1380
Qy 1429 TGCTGTATCTCAAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1488
Db 1381 TGCTGTATCTCAAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Qy 1489 TTAAGAGTCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
Db 1441 TTAAGAGTCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 1549 GGGAACTCTTAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
Db 1501 GGGAACTCTTAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Qy 1609 CACAGGCTTCAAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1668
Db 1561 CACAGGCTTCAAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Qy 1669 CCGGACAGCTTGAAGAGCACTACCTTAGACCTTTGGCTCCCTGAGCAACTGAGCT 1728
Db 1621 CCGGACAGCTTGAAGAGCACTACCTTAGACCTTTGGCTCCCTGAGCAACTGAGCT 1680
Qy 1729 CTTCCATCTCTGAG 1788
Db 1681 CTTCCATCTCTGAG 1740
Qy 1789 TGTCTCCAGAGCTTGAAGCACTTGAAGCACTTGAAGCACTTGAAGCACTTGAAGCACTGA 1848
Db 1741 TGTCTCCAGAGCTTGAAGCACTTGAAGCACTTGAAGCACTTGAAGCACTTGAAGCACTGA 1800
Qy 1849 GGGGCTGT 1908
Db 1801 GGGGCTGT 1860
Qy 1909 AGCGATGT 1968
Db 1861 AGCGATGT 1920
Qy 1969 TGAAGT 2028
Db 1921 TGAAGT 1980
Qy 2029 TGGT 2075
Db 1981 TGGT 2027

Sequence 21, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OR INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 37278
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(37278)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-21

Query Match 66.8%; Score 1938.8; DB 18; Length 37278;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 955 CCGATGCTGTGCTGATCCGATGCTGTTAATGTGAACCCCAACCGTGGCCACAC 1014
Db 25129 CAGATGCTGTGCTGATCCGATGCTGTTAATGTGAACCCCAACCGTGGCCACAC 25188
Qy 1015 TGAAGATGTAGCCAGTCAATGTTGGTCAACTGGGTTACACACCGAGTGGGAGAC 1074
Db 25189 TGAAGATGTAGCCAGTCAATGTTGGTCAACTGGGTTACACACCGAGTGGGAGAC 25248
Qy 1075 AGAAGGCTGCGGAGAGTGGGACCTGTGTGTGCTTTGGCGGGCTCCATGGCGG 1134
Db 25249 AGAAGGCTGCGGAGAGTGGGACCTGTGTGTGCTTTGGCGGGCTCCATGGCGG 25308
Qy 1135 ACTGTACGTCGCTCTCGCGCCCTCTCTGAGAAATGAGACCAAGTGTGACCTTCT 1194
Db 25309 ACTGTACGTCGCTCTCGCGCCCTCTCTGAGAAATGAGACCAAGTGTGACCTTCT 25368
Qy 1195 TCAAGACAGAGTGTGAGAGTGTGAGACCTGATCTGGGCTGAGCGCAACTTCTC 1254
Db 25369 TCAAGACAGAGTGTGAGAGTGTGAGACCTGATCTGGGCTGAGCGCAACTTCTC 25428
Qy 1255 TTAAGAGTCCCGAAGAGAAATGACATGTGCTAAATCTGACAGGTGACCCGGCTGAG 1314
Db 25429 TTAAGAGTCCCGAAGAGAAATGACATGTGCTAAATCTGACAGGTGACCCGGCTGAG 25488
Qy 1315 ATACCTCTTCTGCGCTGCGAAGAGACCTTAAAGTTCGAAAGGCAATTCCTCAAGAAA 1374
Db 25489 ATACCTCTTCTGCGCTGCGAAGAGACCTTAAAGTTCGAAAGGCAATTCCTCAAGAAA 25548
Qy 1375 AGTCTCTACCTGTCTGAGGAGAGTACAGAGACCTTCAAGAACTGACACCGGTGCTG 1434
Db 25549 AGTCTCTACCTGTCTGAGGAGAGTACAGAGACCTTCAAGAACTGACACCGGTGCTG 25608
Qy 1435 ATACTCAAGGAGAGCTGTCTGTGTATCTGTCTGTCTCCAGAGAAAGGCAATCTTAAAG 1494
Db 25609 ATACTCAAGGAGAGCTGTCTGTGTATCTGTCTGTCTCCAGAGAAAGGCAATCTTAAAG 25668
Qy 1495 AGTCTCAAGGAGAGCTGTGTGTATCTGTCTGTCTGTCTCCAGAGACCAAGGAGCTGGGAGAC 1554
Db 25669 AGTCTCAAGGAGAGCTGTGTGTATCTGTCTGTCTGTCTCCAGAGACCAAGGAGCTGGGAGAC 25728
Qy 1555 TCTTGAAGCAGAGT 1614
Db 25729 TCTTGAAGCAGAGT 25788
Qy 1615 CTTCAAGGAGCTCTCTCTCAACCGGCAATTCCTCAAGGAGTGTGTGTGTGTGTGTGTGTGT 1674
Db 25789 CTTCAAGGAGCTCTCTCTCAACCGGCAATTCCTCAAGGAGTGTGTGTGTGTGTGTGTGTGT 25848

QY 1675 CAGCCTTAGAAGGCACTACCCCTTAGACA CTTTGGCTCCCTGGACCACTGGCCCTCTCC 1734
DB 25849 CAGCCTTAGAAGGCACTACCCCTTAGACA CTTTGGCTCCCTGGACCACTGGCCCTCTCC 25908
QY 1735 ATCTGACGCCCGGCCCGACGCCCTCAGGGGCTGTGAGTGAAGACAGCATCTCTGCT 1794
DB 25909 ATCTGACGCCCGGCCCGACGCCCTCAGGGGCTGTGAGTGAAGACAGCATCTCTGCT 25968
QY 1795 CCGAGTCTTTGACCAATTGGAATTGCTGCTGAA CTTTCCGAAACCCCACTGAGGGCT 1854
DB 25969 CCGAGTCTTTGACCAATTGGAATTGCTGCTGAA CTTTCCGAAACCCCACTGAGGGCT 26028
QY 1855 GTGTGTCTGTGACAACCTGAGGGGGCTTGAGGACCTCCCTCGAAGGTCTGAAGCAT 1914
DB 26029 GTGTGTCTGTGACAACCTGAGGGGGCTTGAGGACCTCCCTCGAAGGTCTGAAGCAT 26088
QY 1915 GGTGGCAGGAATCTTGGGGGATAGCTGCTTTCTGACAGACTGCAAGAGTGA CTG 1974
DB 26089 GGTGGCAGGAATCTTGGGGGATAGCTGCTTTCTGACAGACTGCAAGAGTGA CTG 26148
QY 1975 CAGCCTTAGAAGGCACTACCCCTTAGACA CTTTGGCTCCCTGGACCACTGGCCCTCTCC 2034
DB 26149 CAGCCTTAGAAGGCACTACCCCTTAGACA CTTTGGCTCCCTGGACCACTGGCCCTCTCC 26208
QY 2035 CCTAGTATGGGGTATGGCTGAGAGGGTTTGAGAGAACCTGGGCTCGAATTCCTCAG 2094
DB 26209 CCTAGTATGGGGTATGGCTGAGAGGGTTTGAGAGAACCTGGGCTCGAATTCCTCAG 26268
QY 2095 TGAATAGAGTACATCAAGGGCTCTACGCTGACACCTGACTGAACCTGAAAGATGAGAGA 2154
DB 26269 TGAATAGAGTACATCAAGGGCTCTACGCTGACACCTGACTGAACCTGAAAGATGAGAGA 26328
QY 2155 AATGCGATTGATGTGAAAAGGAATGGAA CTTTGTGCTGCCGAGTGTATAGTGGGGTGG 2214
DB 26329 AATGCGATTGATGTGAAAAGGAATGGAA CTTTGTGCTGCCGAGTGTATAGTGGGGTGG 26388
QY 2215 CCTGAAGGTGCTACCTCTTTGTGCGATGAGTGA CCACTGATTTCCCAACCCCTGT 2274
DB 26389 CCTGAAGGTGCTACCTCTTTGTGCGATGAGTGA CCACTGATTTCCCAACCCCTGT 26448
QY 2275 CTCTGGCTGACCTTCAATAGTTTCTGTTCCATCAACCA CCGAGGTTTGAACCTCTGA 2334
DB 26449 CTCTGGCTGACCTTCAATAGTTTCTGTTCCATCAACCA CCGAGGTTTGAACCTCTGA 26508
QY 2335 CTCTGCTGGAGGTATGTGATGATGCTGCAATTA TTTAGAAGAGAAACAGCTCTGTTT 2394
DB 26509 CTCTGCTGGAGGTATGTGATGATGCTGCAATTA TTTAGAAGAGAAACAGCTCTGTTT 26568
QY 2395 CCATCTGTGCTGCTGATCTCAAAAGACCTGGGAGAACTCGA CCGCTGTTTGAACCTCTGA 2454
DB 26569 CCATCTGTGCTGCTGATCTCAAAAGACCTGGGAGAACTCGA CCGCTGTTTGAACCTCTGA 26628
QY 2455 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTA GATCTCAGAGACTTGAACCTTGA 2514
DB 26629 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTA GATCTCAGAGACTTGAACCTTGA 26688
QY 2515 GCTGTTCTTATGACCAAGATGTGATGCTGTGTTCT CAGGCCAACGGGACCTAGA 2574
DB 26689 GCTGTTCTTATGACCAAGATGTGATGCTGTGTTCT CAGGCCAACGGGACCTAGA 26748
QY 2575 ATGTGCTGACTTATTTTATTTTGTGATCTCACTCTG TTTTGGTTTGTGTTTGT 2634
DB 26749 ATGTGCTGACTTATTTTATTTTGTGATCTCACTCTG TTTTGGTTTGTGTTTGT 26808
QY 2635 TGTGTTGTTTGTGTTTATGATGATTTTGTGCTTTCA ATATATGATGATGATGATGCTG 2694
DB 26809 TGTGTTGTTTGTGTTTATGATGATTTTGTGCTTTCA ATATATGATGATGATGATGCTG 26868
QY 2695 GGGAACTCCACTGTGCCACTGAAAGTTTATGTA CAGAGATGATTTGGCAATGATGCTCT 2754
DB 26869 GGGAACTCCACTGTGCCACTGAAAGTTTATGTA CAGAGATGATTTGGCAATGATGCTCT 26928
QY 2755 CTATTCAAGGGGGGTGGGGGCTTTTCAATGATGATG ATCTTGAGACTGTCTGATTTGAG 2814

DB 26929 CTATTCAAGGGGGGTGGGGGCTTTTCAATGATGATG ATCTTGAGACTGTCTGATTTGAG 26988
QY 2815 TCTCAGTCCCTTACACACCCCAAGGCTGGCCACCTCC CTATCTTCACTGTGGCCAAA 2874
DB 26989 TCTCAGTCCCTTACACACCCCAAGGCTGGCCACCTCC CTATCTTCACTGTGGCCAAA 27048
QY 2875 AAAAAAAAAAAAAAAAAAAAAA 2896
DB 27049 AAAAAAAAAAAAAAAAAAAAAA 27070
RESULT 6
US-10-343-514-27
Sequence 27, Application US/10343514
Publication No. US20040132025A1
GENERAL INFORMATION:
APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIORITY APPLICATION NUMBER: PCT/CA01/01109
PRIORITY FILING DATE: 2001-08-02
PRIORITY APPLICATION NUMBER: US 60/222,650
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: US 60/274,613
PRIORITY FILING DATE: 2001-03-12
PRIORITY APPLICATION NUMBER: CA 2,340,780
PRIORITY FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: Patentin version 3.0
SEQ ID NO 27
LENGTH: 2026
TYPE: DNA
ORGANISM: RAT
US-10-343-514-27
Query Match 61.0%; Score 1771.4; DB 18; Length 2026;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 136; Indels 5; Gaps 3;
QY 49 GTGACCTTGAAGCCCGCGCTCAGCGCGCTGCTACTGCTG CCGACCACTCAACTC 108
DB 1 GTGACCTTGAAGCCCGCGCTCAGCGCGCTGCTACTGCTG CCGACCACTCAACTC 60
QY 109 GCGGTCCCGCACCATGAGTGTGGTCTTACTTCCAGCGCC CCGAGCCAGGCTCCTCGG 168
DB 61 GCGGTCCCGCACCATGAGTGTGGTCTTACTTCCAGCGCC CCGAGCCAGGCTCCTCGG 120
QY 169 CTTCCGCGCTGGCTGCGAGAGCGCGCGCTGGCGGACGGG CTCAATCAATGCGCTA 228
DB 121 CTTCCGCGCTGGCTGCGAGAGCGCGCGCTGGCGGACGGG CTCAATCAATGCGCTA 180
QY 229 AACCTCTGATGAAGAACGACGCGGTGAGGCGACCATCA CAACAACAACCTGCGGACC 288
DB 181 AACCTCTGATGAAGAACGACGCGGTGAGGCGACCATCA CAACAACAACCTGCGGACC 240
QY 289 GCTACGAGTCTCTGAGAGCGCTGGCAAGGCACTTACG GGAAGTGAAGAGCACGAG 348
DB 241 GCTACGAGTCTCTGAGAGCGCTGGCAAGGCACTTACG GGAAGTGAAGAGCACGAG 300
QY 349 AGAGCTCGGGGCGCTGGTGGCCATCAAGTCCATCAAGAA GCAAAATCAAAATGAGAGC 408
DB 301 AGAGCTCGGGGCGCTGGTGGCCATCAAGTCCATCAAGAA GCAAAATCAAAATGAGAGC 360
QY 409 AGGATCTGTGCAATACGAGAGGAGATTGAGATCAATGCTT GACTCAACCAACCCACCA 468
DB 361 AGGATCTGTGCAATACGAGAGGAGATTGAGATCAATGCTT GACTCAACCAACCCACCA 420
QY 469 TCATTGCCATCATGAAGTGTGTAAGATAGCAGCAAGATTG TGAATTTGATGAGATGAG 528

Db 421 TCATTGCAATCCATGAGTGTGTAAGAACAGCAGCAAGATTGATGTCATGAGTAGC 480
 Qy 529 CCAGCCGAGGAGATCTGTATGATTAATCATAGTAGGCGGCCAAGCTGAGTAGAGGAGAGC 588
 Db 481 CCAGCCGAGGAGATCTGTATGATTAATCATAGTAGGCGGCCAAGCTGAGTAGAGGAGAGC 540
 Qy 589 CCAGGCAATTTCTTCGACAGATCTGTGCTGCTGCACTAGTGCACAGAACGGAGATG 648
 Db 541 CCAGGCAATTTCTTCGACAGATCTGTGCTGCTGCACTAGTGCACAGAACGGAGATG 600
 Qy 649 TTACCCAGATCTACAGTGGAAAAATCTTTAGATGCAATGAAACATCAAGATTG 708
 Db 601 TTACCCGAGATCTACAGTGGAAAAATCTTTAGATGCAATGAAACATCAAGATTG 660
 Qy 709 CTGACTTTGGCTCTCCAACTGTACCAAGGCAAGTTCTCCAGACGTTCTGTGGGA 768
 Db 661 CTGACTTTGGCTCTCCAACTGTATACCAAGGCAAGTTCTCCAGACGTTCTGTGGGA 720
 Qy 769 GCCCTCTCTACGCTCGCTGAGATGATCAACGGAAGCCCTATGTGCGCCAGAGGTG 828
 Db 721 GCCCTCTCTATGCTGCTGAGATGATCAACGGAAGCCCTATGTGCGCCAGAGGTG 780
 Qy 829 ACAGCTGTCTCTGAGGCTTCTCTGTATCATCTGTGTGATGCAATGCTCTTTGACG 888
 Db 781 ACAGCTGTCTCTGAGGCTTCTCTGTATCATCTGTGTGATGCAATGCTCTTTGACG 840
 Qy 889 GGAAGATCAATPAAACCTGCTGAGAAATCAATACGAGGCTTACCTGAGGCGGCCA 948
 Db 841 GGAAGATCAATPAAACCTGCTGAGAAATCAATACGAGGCTTACCTGAGGCGGCCA 900
 Qy 949 AGCCGCTCGATGCTGTGAGCTGATCCGATGCTGTATATGATGAACCCCAAGCTGAG 1008
 Db 901 AGCCGCTCGATGCTGTGAGCTGATCCGATGCTGTATATGATGAACCCCAAGCTGAG 960
 Qy 1009 CCAACTGAGAGATGTAACCAATGATGATGATGATCAATGAGGATTAACCAAGGAGTGC 1068
 Db 961 CCAACTGAGAGATGTAACCAATGATGATGATGATCAATGAGGATTAACCAAGGAGTGC 1020
 Qy 1069 GGAAGACGGAAGACCTGCTGAGAGGTGAGGACCTTAATGATGATTTGCGCGGCTTCA 1128
 Db 1021 GGAAGACGGAAGACCTGCTGAGAGGTGAGGACCTTAATGATGATTTGCGCGGCTTCA 1080
 Qy 1129 TGCGGAGCTGATGATGCTGCTGCTGCGGCCCTCTCTGAGATATGAGCCAAAGTGTGA 1188
 Db 1081 TGCGGAGCTGATGATGCTGCTGCTGCGGCCCTCTCTGAGATATGAGCCAAAGTGTGA 1140
 Qy 1189 GCTTCTTCAAGCAGCAGCTGCGGAGGTGAGACACTGTATCTGTGGCTGAGCGGCAAC 1248
 Db 1141 GCTTCTTCAAGCAGCAGCTGCGGAGGTGAGACACTGTATCTGTGGCTGAGCGGCAAC 1200
 Qy 1249 ATTCTCTTAAGAGATCCCGAAAGAAATGATATGCTCAAAATCTGCAAGATACCCGG 1308
 Db 1201 ATTCTCTTAAGAGATCCCGAAAGAAATGATATGCTCAAAATCTGCAAGATACCCGG 1260
 Qy 1309 CTGAGATACCTCTTCTGCGCTGCGCAAGAGCAGCTTAAGCTTCGAAAGGCAATTCTCA 1368
 Db 1261 CTGAGATACCTCTTCTGCGCTGCGCAAGAGCAGCTTCGAAAGGCAATTCTCTCA 1320
 Qy 1369 AGAAAAAGTCTTACTCTGCTGAGGAGGTATCAGAGAGACCTTCAGAGAACTCAGACCG 1428
 Db 1321 AGAAAAAGTCTTACTCTGCTGAGGAGGTATCAGAGAGACCTTCAGAGAACTCAGACCG 1380
 Qy 1429 TGCTGTATCTCCAGGAGAGCTGTCTCTGTATCTCTGCTGCTCCAAAGAAAGCAATCC 1488
 Db 1381 TGCTGTATCTCCAGGAGAGCTGTCTCTGTATCTCTGCTGCTCCAAAGAAAGCAATCC 1440
 Qy 1489 TTAAAGATCTCAGAGAGTAAATCTGATTAATCTCTCTCAGAGCCAGAGTCTG 1548
 Db 1441 TTAAAGATCTCAGAGAGTAAATCTGATTAATCTCTCTCAGAGCCAGAGTCTG 1500
 Qy 1549 GGAAGCTCTTAAGAGCAGTATGTTTGTGATGAGGAGACCCGTGAGAGCAAGTCTC 1608

Db 1501 GGAAGCTCTTAAGACAGCAGTATGTGTTGTGATGAGGAGACCCCGTAGACAGAAATCTC 1560
 Qy 1609 CACAGGCTTACGAGGCTCTCTCTCCACGCAAGGCAATTTCAATCTAATGGCAAGTTCT 1668
 Db 1561 CACAGGCTTACGAGG---CGCTCCATGCAAGGAGCATCTCAAACTCAATGGCAAGTTT 1617
 Qy 1669 CCGGACAGGCTTAAAGAGCACTACCTCTAGCATTTGGCTTCTGTAGCAACTGGCTT 1728
 Db 1618 CCGGACAGGCTTAAAGAGCACTGCCCCCTAGCATCTTTGGCTTCTGTAGCAACTGGCTT 1677
 Qy 1729 CTTCCCATCTGAGACCCCGGACCGCCCTCAGGGGCTGTGAGTGAAGAGACAGATCC 1788
 Db 1678 CTTCCCATCTTACAGCCCGGACCGCTTCTCGAGAGCTGTGAGTGAAGAGACAGATCC 1737
 Qy 1789 TGTCTTCGAGCTCTTGAACCAATGATGCTGTGACATGCTTCCGAAACCCCACTGA 1848
 Db 1738 TGTCTTCGAGCTCTTGAACCAATGATGCTGTGACATGCTTCCGAAACCCCACTGA 1797
 Qy 1849 GGGGCTGTGTCTGTGAGCAACCTGAGGAGGCTTGAAGAGCTCTCTCAGAGTCTGA 1908
 Db 1798 GGAAGCTGTGTCTGTGAGCAACCTGAGGAGGCTTGAAGAGCTCTCTCAGAGGCTTAA 1857
 Qy 1909 AGGATGATGAGCAAGATCTTGGGAGATGCTGCTTTCTGTGACAGACTGCGCAAGAG 1968
 Db 1858 AGGATGATGAGCAAGATCTTGGGAGATGCTGCTTTCTGTGACAGACTGCGCAAGAG 1917
 Qy 1969 TGACTGAGCTTACAGACAGCCCTAGAGATGCTCTCAAGCTCAGCTGAGAGGAGAGA 2028
 Db 1918 TGACAGAGCTTACAGACAGCCCTAGAGATGCTCTCAAGCTCAGCTGAGAGGAGAGA 1977
 Qy 2029 TGTGCTCTAGT-ATGAGGATGAGCTGTGAG-GGGTTTGAAGAGAAC 2075
 Db 1978 CAGTGCCCAAGTATGAGGATGAGCTTAAAGAGGCTTTGAGAGAGAAC 2026

RESULT 7
 US-09-963-159-1
 ; Sequence 1, Application US/09963159
 ; Patent No. US2002007712A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CURTIS, RORY A.J.
 ; APPLICANT: GALVIN, KATHERINE M.
 ; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
 ; FILE REFERENCE: 10147-5001
 ; CURRENT APPLICATION NUMBER: US/09/963,159
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 60/234,922
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 3353
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (3268)
 ; NAME/KEY: unsure
 ; LOCATION: (3270)
 ; NAME/KEY: unsure
 ; LOCATION: (3272)
 US-09-963-159-1

Query Match 47.5%; Score 1379; DB 9; Length 3353;
 Best Local Similarity 81.8%; Pred. No. 0;
 Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;
 Qy 82 CTACTGTGCGCCGACCACTTCACCTGCGGATCCCGCAGACATGAGTGGTGGCTTAC 141
 Db 116 CTACTGATTTCCCGCGCCCTTGTCTACCTCTGCTGCGCATGAGATCCCTGTTTTCG 175
 Qy 142 TCAGCGCCGAGCAGGCTTCCCTGCGCTTCCGCTGCGCTGAGAGAGCGCCGCGCG 201

Db 176 CGCGGCGCTCCGGCCCCCACTCCCTCGGCGCGAGAGCTA-----GCCCGGCGCG 223
QY 202 TGGCGGACGGGCTCATCAAGTCGCTTAACTCTGATGAGAGACAGCGGTGAAAGCGC 261
Db 224 TGGGGGAGGGGCTGATCAAGTCGCGCAAGCCCTTAAGAAAGACAGCGGTGAAAGCGC 283
QY 262 ACCATCAAAACAAACCTCGGCGACCGCTTCAAGTTCTTGAGAGACGCTGGGCAAGGCA 321
Db 284 ACCACCAAGGACAAACCTCGGCGACCGCTTCAAGTTCTTGAGAGACGCTGGGCAAGGCA 343
QY 322 CTTACGGGAAAGTGAAG 381
Db 344 CTTACGGGAAAGTGAAG 403
QY 382 TCAGAGAAAGCAAAATCAAGATGAGAGAGATCTGCTGACATACGAGAGAGAGATTAGA 441
Db 404 TCGGAAAGGACAAATCAAGATGAGAGAGATCTGATGACATCGAGAGAGAGATTAGA 463
QY 442 TCAGATCTTCACTCAACACACCCCAATCATTCATTCATTCATTCATTCATTCATTCAT 501
Db 464 TCATGTCATCATCAACACACCCCAATCATTCATTCATTCATTCATTCATTCATTCAT 523
QY 502 GCAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 524 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
QY 562 AGCGGCGACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
Db 584 AGCGGCGACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 643
QY 622 TGCACTACTGCGACACAGAGCGGAGTCGTTACGAGATCTCAAGCTGAGAAACATCTTC 681
Db 644 TGCACTACTGCGACACAGAGCGGAGTCGTTACGAGATCTCAAGCTGAGAAACATCTTC 703
QY 682 TAGATGCAATGGAACATCAAGATGCTGATCTTGGCTCTCCAACTGATGACCAAG 741
Db 704 TGATGCAATGGAATATCAAGATGCTGATCTTGGCTCTCCAACTGATGACCAAG 763
QY 742 GCAAGTCTCTCAGAGCTTCTGTTGGAGCCCTCTCTACGCTCGCTGAGATGATCAAG 801
Db 764 GCAAGTCTCTCAGAGCTTCTGTTGGAGCCCTCTCTACGCTCGCTGAGATGATCAAG 823
QY 802 GGAAGCCCTATGTTGGGCGCAGAGGTGACAGCTGCTGCTGGGGCTTCTCTGATCAATCC 861
Db 824 GGAAGCCCTATCAGGCGCAGAGGTGACAGCTGCTGCTGGGGCTTCTCTGATCAATCC 883
QY 862 TGGTGCATGAGCACCATGCTTTTGAACGGGCGAGATCAATTAACATGCTGAGAGAAATCA 921
Db 884 TGGTGCATGAGCACCATGCTTTTGAACGGGCGAGATCAATTAACATGCTGAGAGAAATCA 943
QY 922 GTTACGGGGCTTACCGTGAAGCGGCCCAAGCCGTCGATGCTGTGGCTGATCCGTTGCG 981
Db 944 GCAACGGGGCTTACCGGAGGACCACTTAACCCCTCTGATGCTGTGGCTGATCCGTTGCG 1003
QY 982 TGTTAATGAGAACCCCAACCCCTCGGGCGACACTGAGAGATGATGACCAATGATTTGGGG 1041
Db 1004 TGTGATGAGTAACCCCAACCCCTCGGGCGACACTGAGAGATGATGACCAATGATTTGGGG 1063
QY 1042 TCAACTGGGGTTACACACCGGAGTCCGGGAAACAGAAAGCCCTGAGAGGAGTGGACCC 1101
Db 1064 TCAACTGGGGCTACACACCGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123
QY 1102 CTAGTGTGACTTTGGCGGGGCTCATGCGGAGCTGTTACGTCCTCTCGCGCCGCC 1161
Db 1124 CTGGCAGTGACTGCGCGCGCTCATGCGGAGCTGTTACGTCCTCTCGCGCCGCC 1183
QY 1162 TCCGAGAGATGAGACCAAGTGTGAGCTTTTCAAGCAGACAGTGCCTGGAGAGTGA 1221
Db 1184 TCCGAGAGATGAGACCAAGTGTGAGCTTTTCAAGCAGACAGTGCCTGGAGAGTGA 1243
QY 1222 GCACTGATCTGGGCTGAGCGGCAATTTCTTAAGAGTCCGAAAGAGAGATGACA 1281
Db 1244 GCACACCCCTGGCTGGAGCGCCAGCATTCGCTCAAGAAAGTCCCGCAAGAGATGACA 1303

QY 1282 TGGCTCAAAATCTGCAAGTGAACCCGGCTGAGAGATACCTTTCTCGCCCTGGCAAGCA 1341
Db 1304 TGGCCCACTCTCCACAGAGACAGGCTGATGACATGCTCCATCGCTTGGCAAGCA 1363
QY 1342 GCGTTAAGCTTCCGAAAGGCACTTCAAGAAAGTCTCTTCACTCGTCAAGGAGGATC 1401
Db 1364 ACCTCAAGCTGCCAAGAGGATTTCTCAAGAAAGTGTGATGCTTCTGCAAGAGGATC 1423
QY 1402 AGGAGACCTCTCAGAACTCAGACCGAGTGTGATATCTCAGAGGACGCTGCTGCTG 1461
Db 1424 AGGAGACCTCTCAGAGCTCAGCCCAATCCTCGAGCCCAAGGAGGCTGCTGCTGCT 1477
QY 1462 TATCCTGCTCCCAAGAAAGGCACTTCAAGAGTGTGATGATGATGATGATGATGAT 1521
Db 1478 ---CCCTGCTCCCAAGAAAGGCACTTCAAGAAAGGCTCAGAGGAGGATGCTGCTACT 1534
QY 1522 ACTCTCTCCAGAGCCCAAGAGTCTGGGGAATCTTTAGACGCGCAGTATGATTTGTGA 1581
Db 1535 ACTCTCTCCAGAGCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1594
QY 1582 GTGGGACCCCGTGGAGCAGAACTTCAAGGCTTCAAGGCTCTCTTCCACCGCAAG 1641
Db 1595 GTGGGATCCCAAGAGCAGAACTTCCAGGCTTCAAGGCTCTCTTCCACCGCAAG 1654
QY 1642 GCATTTCAAACTCAATGAGATTTCTCCGCAACGCTTTAAGAGGATCACTCCCTAAGCA 1701
Db 1655 GCATTTCAAACTCAATGAGATTTCTCCAGACAGCTTTAAGAGCTGCGGCCCCCA 1714
QY 1702 CTTTGGCTCCCTGAGCAACATGGGCTCTCCATCCATCTGAGGCGCGGCGCGGCGCT 1761
Db 1715 CTTTGGCTCCCTGAGTGAATCTGCGCCACCTGCGCCCTGCGGCGCGGCGCGGCGCT 1774
QY 1762 CAGGGCTGTGATGAGCAGCATCTGCTCTCCAGATCTTTGAACCAATGAGATTCG 1821
Db 1775 CAGGGCTGTGAGCAGCATCTGCTCTCCAGATCTTTGAACCAATGAGATTCG 1834
QY 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1881
Db 1835 CTGAAGCTCTTCCGAAACCCCACTGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1894
QY 1882 TTGAGCAGCTCTCCTAGAG-----GTCTGAAGGAGATGCTGAGAAATCTT 1929
Db 1895 TTGAGAGCCCTCCTAGAGGAGGCTGAGAGCTGAGAGGCTGAGAGGCTGAGAGGCTT 1954
QY 1930 TGGGAGATGAGCTTTTCTGACAGACTGCCAAGAGTGAAGTGAAGCTTACAGACAAG 1989
Db 1955 TGGGAGCAGCTGCTTTTCTGACAGACTGCCAAGAGTGAAGTGAAGCTTACAGACAAG 2014
QY 1990 CCTTGAAGATCTGCTCAAGCTCAGCTGAGAGAGAGATGATGATGATGATGATGATG 2049
Db 2015 CACTGAGGCTGCTCAAGAGCTCAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2073
QY 2050 GCTCTGAGAGGTTTGCAGAGAACTTGGTGGGATTCCTCC 2092
Db 2074 GCTCTGAGATGAGCTGTTGACCCCGAAGGAGATGCTTCT 2116

RESULT 8
US-10-423-543-43
; Sequence 43, Application US/10423543
; Publication No. US2004058355A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Lieberman, Rosana K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Myoung

APPLICANT: Williamson, Mark J.
APPLICANT: Silos-Santiago, Inmaculada
APPLICANT: Bhandari, Rajasekhari
TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MP103-0230NMIM
CURRENT APPLICATION NUMBER: US/10/423, 543
CURRENT FILING DATE: 2003-04-25
PRIORITY APPLICATION NUMBER: US 10/278, 036
PRIORITY FILING DATE: 2002-10-22
PRIORITY APPLICATION NUMBER: US 09/711, 216
PRIORITY FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: US 60/205, 447
PRIORITY FILING DATE: 2000-05-19
PRIORITY APPLICATION NUMBER: US 10/012, 055
PRIORITY FILING DATE: 2001-11-13
PRIORITY APPLICATION NUMBER: US 60/248, 325
PRIORITY FILING DATE: 2000-11-14
PRIORITY APPLICATION NUMBER: US 10/003, 690
PRIORITY FILING DATE: 2001-11-15
PRIORITY APPLICATION NUMBER: US 60/248, 893
PRIORITY FILING DATE: 2000-11-15
PRIORITY APPLICATION NUMBER: US 09/797, 039
PRIORITY FILING DATE: 2001-02-28
PRIORITY APPLICATION NUMBER: US 60/186, 061
PRIORITY FILING DATE: 2000-02-29
PRIORITY APPLICATION NUMBER: US 10/217, 168
PRIORITY FILING DATE: 2002-08-12
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 119
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 3353
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(3353)
OTHER INFORMATION: n = A,T,C or G
US-10-423-543-43

Query Match 47.5%; Score 1379; DB 17; Length 3353;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;

QY 82 CTACTGTCGCCGACCACTTCACCTCGCGGTCCCGCACCAGTGGAGTGGCTTAC 141
DB 116 CTAAGTATTCCTCCCGCGCTTCTCACTCGCTGCTGCAATGAGTGGCTTTCG 175
QY 142 TCCAGCGCCGAGCGCAAGTCTCTCGCGCTCCCGCTCGAGAGCGCCGCGCGC 201
DB 176 CGCGCGCTCCCGCGCGCACTCTCGCGCGCAAGTCA-----GCCGCGCGC 223
QY 202 TGGGGAAGGAGTCAAGTGGCTAAACCTCGATGAGAGAGAGAGAGAGAGAGAG 261
DB 224 TGGGGAAGGAGTCAAGTGGCTAAACCTCGATGAGAGAGAGAGAGAGAGAGAG 283
QY 262 ACCATCAAAACAACTCGCGGACCGCTTCAAGTCTTGGAGAGCGCTGGGCAAGGCA 321
DB 284 ACCACCAAGAGCAACCTCGCGGACCGCTTCAAGTCTTGGAGAGCGCTGGGCAAGGCA 343
QY 322 CTAAGGGAAGTGAAG 381
DB 344 CTAAGGGAAGTGAAG 403
QY 382 TCAGGAAGAGCAAAATCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441
DB 404 TCAGGAAGAGCAAAATCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
QY 442 TCAGTCTTCACTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501

DB 464 TCATGTCATCACTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 523
QY 502 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
DB 524 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 583
QY 562 AGCGGCAAG 621
DB 584 AGCGGCAAG 643
QY 622 TGCATGTCATCACTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 681
DB 644 TGCATGTCATCACTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 703
QY 682 TAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 741
DB 704 TAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 763
QY 742 GCAAGTCTTCCAG 801
DB 764 GCAAGTCTTCCAG 823
QY 802 GGAAGCCCTATGAG 861
DB 824 GGAAGCCCTATGAG 883
QY 862 TGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 921
DB 884 TGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 943
QY 922 GTAAAGGAGGCTTACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
DB 944 GCAAGGAGGCTTACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1003
QY 982 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041
DB 1004 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1063
QY 1042 TCAATGAGGAGTCAACAG 1101
DB 1064 TCAATGAGGAGTCAACAG 1123
QY 1102 CTAAGTATTCCTCCCGCGCTTCTCACTCGCTGCTGCAATGAGTGGCTTTCG 1161
DB 1124 CTAAGTATTCCTCCCGCGCTTCTCACTCGCTGCTGCAATGAGTGGCTTTCG 1183
QY 1162 TCTGAGAGATGAG 1221
DB 1184 TCTGAGAGATGAG 1243
QY 1222 GCACTGATCTTGGGCTGAG 1281
DB 1244 GCACTGATCTTGGGCTGAG 1303
QY 1282 TGGGCAAAATGCAAGTGAAG 1341
DB 1304 TGGGCAAAATGCAAGTGAAG 1363
QY 1342 GCTTAAAGCTTCCGAAAG 1401
DB 1364 ACCTCAAGTGCAG 1423
QY 1402 AG 1461
DB 1424 AG 1477
QY 1462 TATCCCTGCTCCGAAAG 1521
DB 1478 ---CCCTGCTCCGAAAG 1534
QY 1522 ACTCTCTTCCAG 1581
DB 1535 ACTCTCTTCCAG 1594

QY	1582	CTGGGGGACCCCGCTGGAGCAGAAAGTCTCCACAGGCTTCAGGGCTCTCTCTCCACGGCAAG	1641
Db	1595	GTGGGATCTCCAAAGAGCAGAAAGCCCTCGCAAGCTTCAGGGCTCTCTCTCCATTCGCAAG	1654
QY	1642	GCAATCTGAATCTCAATGAGCAAGTTCTCCCGACAGACCTTAGAAGGCACTACCCCTAGCA	1701
Db	1655	GCATCCTCAAACTCAATGAGCAAGTTCTCCCAACAGACCTTAGAAGCTCCGGGCCCAACA	1714
QY	1702	CCTTTGGCTCCCTGGACCAACTGGCTCTCCCAATCTGGACGCCGGGCCCAAGCCCTCT	1761
Db	1715	CCTTGGCTCCCTGGATGAATCGGCCCACTCGGCCCTCCGGCCCGGGGCAGCCGACCT	1774
QY	1762	CAGGGGCTGTGAGTGAAGGACAGCATCTGTCTCCGAGTCCCTTTAGCAATTTGACTTGG	1821
Db	1775	CAGGGGCTGTGAGCAGAGCAGACATCTGTCTCTGAGTCTTTAGCAAGCTTGAATTGG	1834
QY	1822	CTGAACGTCTTCCCGAAACCCCACTGAAGGGGCTGTGTCTGTGCAACAACCTGAAGGGGC	1881
Db	1835	CTGAACGGCTCCCAAGAGCCCACTGCGGGGCTGTGTGTGTGCAACAACCTCAGGGGC	1894
QY	1882	TTGAGCAGCTCTCCCTCCAGAG-----GTCTGAAGCGATGTGTGGCAGGAATCTT	1929
Db	1895	TTGAAGAACCCCTCCAGAGGGCCCTGGAAGCTGTGCTGAGGGCTGCGGCGAGGATCCTT	1954
QY	1930	TGGGGGATAGCTGCTTTTCTGTGACAGACTGCCAAGAGGTGACATGCACTTACAGCAAG	1988
Db	1955	TGGGGGACAGCTGCTTTTCTGTGACAGACTGCCAAGAGGTGACAGCACTTACCGACAG	2014
QY	1990	CCTTAGAATCTGCTCAAAAGCTCAGCTGAGAGGAAGAGATGTGTCCCTAGTATGGGGTAG	2049
Db	2015	CAGTGAAGGTGTGCTCAAAAGCTCAGCTGAGTGAATGAGCAATGGCCAG--CCGGTAG	2073
QY	2050	GCTCGAAGAGGTTTGGCAGAGAAACCTTGGGTGGATTTCTTC	2092
Db	2074	GCTCGAATGCAAGCTGTGTGACCCCGAGGGGAGATGCTTTC	2116
RESULT 9			
US-10-322-281-25			
; Sequence 25, Application US/10322281			
; Publication No. US20040126762A1			
; GENERAL INFORMATION:			
; APPLICANT: David W. Morris			
; APPLICANT: David S. Malandro			
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer			
; FILE REFERENCE: 529452001000			
; CURRENT APPLICATION NUMBER: US/10/322.281			
; NUMBER OF SEQ ID NOS: 866			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 25			
; LENGTH: 3404			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-322-281-25			

Query Match	47.4%	Score 1375.8	DB 18	Length 3404
Best Local Similarity	81.7%	Pred. No. 0		
Matches 1652	Conservative	0	Mismatches 337	Indels 34
			Gaps	4
QY	82 CTACTGCTGCCCGACCACTCACTCGCGGTCCCCGACCATGAGTCGGTGGCTTAC	141		
Db	87 CTACTGATTCCTCCCTGCGCCGCTTGCTCACTTCCTGCGCATGAGTCGGTGTTCG	146		
QY	142 TCCAGCGCCGAGCCAGGCTCCCTGGGCTCGGCCCTGGAGAGCCCCGCGCCG	201		
Db	147 CGCGCGCTCCGCGCCCACTCTCTCGCGCGAGACTA-----GCCCGCGCG	194		
QY	202 TGGCGAGAGGGCTCATCAAGTCGCTTAACCTGATAGAAGCAGCGGTGAAGCGGC	261		
Db	195 TGGCGAGAGGGCTATCAAGTCGCCCAAGCCCTTAATAGAAGCAGCGGTGAAGCGGC	254		

QY	262	ACCATCAAAACAAACCTGGGGCACCGCTTACGAATTCCTGGAGACGCTGGGCAAGGCA	321
Db	255	ACCAACAACAAGCAACCTGGGGCACCGCTTACGAATTCCTGGAGACCTGGGCAAGGCA	314
QY	322	CCTACGGGAAGGTGAAGAAAGGCACAGAGAGCTCGGGGCGTCTGTGGCCATCAAGTCCA	381
Db	315	CCTACGGGAAGGTGAAGAAAGGCGCGGAGAGCTCGGGGCGCTGTGGCCATCAAGTCCA	374
QY	382	TCAGAAAGAACAATTCAGAGATGACAGAGATCTGCTGCACATACGAGAGGAGATTGAGA	441
Db	375	TCGGAGAGACAATAATCAAAAGATGAGCAAGATCTGATGCACATACGAGAGGAGATTGAGA	434
QY	442	TCATGTCCTCACTCAACCAACCCCAATCATTTGCCATCATAGAAATGTTTGAATPAGA	501
Db	435	TCATGTCATCACTCAACCAACCTTCAATCATTTGCCATCATAGAAATGTTTGAACAGAGA	494
QY	502	GCAAGATTGTGATTTGATTCATGAGATATGTCACAGCCGAGGCGATCTGATGATTACATAGT	561
Db	495	GCAAGATGTGATTCGTCATGAGATATGTCACAGCCGAGGCGACTTTATGACTACATCAGCG	554
QY	552	AGCGGCCACGCGTGAAGTGAAGCGGAGCGCACGAGCATTTTCTTCGACAGATCGTGTCTGCC	621
Db	555	AGCGGACAGAGCTCAGTAGAGCGGAGAGCTAAGGCAATTTCTTCGAGATTCGTCGCG	614
QY	622	TGCATTACTGCCACCAAGACGGGATCGTTCAACGAGATCTCAGCTGAGAAAATCTCTTC	681
Db	615	TGCATTAATTCGATCAGAAACAGAGTTGTCACCGGATCTCAGCTGAGAAAATCTCTCT	674
QY	682	TAGATGCGCAATGGAACACATCAAGATGTGTGACTTTGGGCTCTCCACCGTATCCAAG	741
Db	675	TGATATCCAAATGGAAATATCAAGATGTGACTTTGGGCTCTCCAACTCTACACATCAAG	734
QY	742	GCAAGTTCCTCAGACGTTCTGTGGGAGCCCTCTCTACGCTCGGCTGAGATGTCAG	801
Db	735	GCAAGTTCCTGAGACATTTGTGGGAGCCCTCTCTATGCTCGCAGAGATTTGTCAGT	794
QY	802	GGAAGCCCTAATGTGGGCCCAAGAGGTGACAAGTGTCTTGGGCGCTTCTCTGATCATCC	861
Db	795	GGAAGCCCTTACAAGGCCCAAGAGGTGACAAGTGTCTCTGGGATGTTCTCTCTCATCC	854
QY	862	TGGTGCATGGGACCAATGCGCTTTGACGGGCAAGATCATAAACACTGTGTGAAGCAATCA	921
Db	855	TGGTGCATGGGACCAATGCGCTTTGATGGGCAATGACATTAAGATCTTAATGAACAGATCA	914
QY	922	GTAACGGGGCTTACCGGTAGGCGGCCCAAGCGGTCCGATGCGCTGTGGGCTGTATCCGGTAC	981
Db	915	GCAACGGGGCTTACCGGGAGGCCACTTAACCTCTGATGCTCTGGGCTGTATCCGGTAC	974
QY	982	TGTTAATGTGAACCCCAACCCGTGGGGCCACAATGAGAGATGTACCAGTCAATTTGTGGG	1041
Db	975	TGTTGATGTGAACCCCAACCCGCGGGCCACCTGAGAGATGTGGCCAGTCACTGTGGG	1033
QY	1042	TCAACTGGGGGTTCACACACCGGAGTCGGGGAAACAGAAAGCCCTGTGTAGGGTGGCACCC	1101
Db	1035	TCAACTGGGGGTTCACGCCCAACCGAGTGGGAGACAGAGAGCTTCGATAGGGTGGGCACC	1093
QY	1102	CTAAGTGTGACTTTGGGCGGGGCTTCATATGGGAGACTGGTTTACGTCAGTCCGCGGCCCC	1161
Db	1095	CTGCGAGTGAATCTGCGCGGCTTCATATGGCTGATGGCTCTCGGGGTTCCTCCGCCCCC	1155
QY	1162	TCTTGAGAAATGAGGCCAAGGTGTGACAGCTTCTTCAAGCAGCAGTGCCTGGGAGGTGAGA	1221
Db	1155	TCTTGAGAAATGAGGGGCCAAGGTGTGACAGCTTCTTCAAGCAGCAGTGCCTGGTGGGGAA	1214
QY	1222	GCACTGTACTTGGGCTGGAGCGGCAATTTCTTTAAGAGTCCCGAAAGGAATGACA	1281
Db	1215	GCACCAACCCCTGGGCTGGAGCGGCAATTTGCTTCAAGAAATCCCGCAAGAGGAATGACA	1273
QY	1282	TGGCTCAAAATCTGCAAGGTGACCCGCGTGAAGATACCTCTTTCGCCCTGTGGCAAGACA	1341
Db	1275	TGGCCCAATCTCTCCACGTGACAACGCGTGTATGACATGCCCATATGCCCTGTGGCAAGACA	1334
QY	1342	GCCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAGTCCCTTACTGCTGTCAAGGGAGGTAC	1401


```

Db      1335  ACCCTGAGCTGCCAAGGAGGATCTCAAGAGAGAGGTCAGCCTCTGCAAGAGGGGTAC 1394
Qy      1402  AGGAGAGCCCTCAGAGACTCAGACCGGNGCTGTACTCCAGGGGAGCTGTCTCTGTG 1461
Db      1395  AGAGAGACCTTCGAGACTCAGCCCAATCCTTCGAGCCCAAGGAGGAGCTGCC----- 1448
Qy      1462  TATCCCTGCTCCAGAGAAAGGCACTCTTAAGAGTCTGACAGCTGAACTTGCTTACT 1521
Db      1449  ---CCCTGCTCCCAAGAGGAGCTTCTCAAGAGGCCCAAGCGGAGTCTGTGCTACT 1505
Qy      1522  ACTCTCTCCAGAGCCCAAGAGTCTGAGGAACTTTAGAGCCAGAGTGTGTGTGA 1581
Db      1506  ACTCTCTCCAGAGCCCAAGAGTCTGAGGAGCTTTGAGAGCGAGGAGCGTGTGTGA 1565
Qy      1582  GTGGGAGCCCGAGAGAGAGTCTCAGAGCTTCAAGGCTCTCTCCACAGCGAGG 1641
Db      1566  GTGGGAGTCCCAAGAGAGAGAGCTTCAGAGCTTCAAGGCTCTCTCTCAATGCAAG 1625
Qy      1642  GCATTCCTCAACTCAATGAGAGTCTCCGCAAGCTTGAAGGCACTAACCTTACGA 1701
Db      1626  GCATTCCTCAACTCAATGAGAGTCTCCAGAGAGCTTGAAGCTTGGGCGCCCAACA 1685
Qy      1702  CTTTGGCTCTCCAGAGAGTCTGCTCTCCATCTGAGAGCCGCGCCCAAGCCCTT 1761
Db      1686  CTTTGGCTCTCCAGAGAGTCTGCTCTCCATCTGAGAGCCGCGCCCAAGCCCTT 1745
Qy      1762  CAGGGGCTGTGAGAGAGAGAGTCTGCTCTCCAGAGCTTGAAGCAATGAGCTTGC 1821
Db      1746  CAGGGGCTGTGAGAGAGAGAGTCTGCTCTGAGTCTTGAAGCAATGAGCTTGC 1805
Qy      1822  CTGAACCTCTTCCAGAGAGAGAGTCTGAGGAGCTGTGTCTGAGAGCACTGAGGAGG 1881
Db      1806  CTGAACGAGCTTCCAGAGAGAGAGTCTGAGGAGCTGTGTCTGAGAGCACTGAGGAGG 1865
Qy      1882  TTGAGAGAGCTCTCCAGAG-----GTCTGAAGCAATGAGTGGAGAGAGTCTT 1929
Db      1866  TTGAGAGAGCTCTCCAGAGAGAGAGTCTGAGGAGCTGTGTCTGAGAGAGTCTT 1925
Qy      1930  TGGGGGATGAGTCTTCTTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 1989
Db      1926  TGGGGGATGAGTCTTCTTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 1985
Qy      1990  CCTGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 2049
Db      1986  CACTGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 2044
Qy      2050  GCTCTGAGAGGTTTTCAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 2092
Db      2045  GCTCTGAGAGGTTTTCAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTT 2087

```

```

RESULT 10
US-10-370-715B-639
; Sequence 639, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
;   Patin Docket Preview
;   APPLICANT: BODARY, SARAH C.
;   APPLICANT: CLARK, HILARY
;   APPLICANT: BRISDELL, HUNTE
;   APPLICANT: JACKMAN, JANET
;   APPLICANT: SCHOENFELD, JILL R.
;   APPLICANT: WILLIAMS, P. MICKEY
;   APPLICANT: WOOD, WILLIAM I.
;   APPLICANT: WU, THOMAS D.
;   TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
;   TITLE OF INVENTION: Related Diseases
;   FILE REFERENCE: P19481-US
;   CURRENT APPLICATION NUMBER: US/10/370,715B
;   CURRENT FILING DATE: 2003-02-21
;   NUMBER OF SEQ ID NOS: 742
;   SEQ ID NO 639

```

```

; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-370-715B-639
Query Match      47.4%; Score 1374.2; DB 18; Length 3443;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy      82  CTACTGTGCGCCAGCCACTTCACCTGCGGATCCCGCCCAAGAGAGTCTGAGCTTAC 141
Db      87  CTACTGTATTCCTCCGCGCCCTTCTGCTCACTCTGCTGCGCATGAGATCTGCTTTTCG 146
Qy      142  TCCAGCGCCGAGCCAGAGCTCTCTGCGCTCTGCGCTCTGCGCTCTGAGAGCGCCGCGC 201
Db      147  CGGCGCGCTCCGCGCCCACTCTCTGCGCCGAGAGCTA-----GCCGCGCGC 194
Qy      202  TGGCGAGCGGCTCATTAAGTCTGCTTAACCTCTGATGAGAGAGAGAGAGAGAGGCGC 261
Db      195  TGGCGAGAGGCTGATCAATGCTGCGCAAGCCCTTAATGAAGAGAGAGAGAGAGAGGCGC 254
Qy      262  ACCATCAAAACAACAACCTGCGGCAACGCTTACAGAGTCTTGAAGAGCTGAGGCAAGGCA 321
Db      255  ACCATCAAAACAACAACCTGCGGCAACGCTTACAGAGTCTTGAAGAGCTTGAAGGCA 314
Qy      322  CTTACGGGAGAGTGAAGAGAGAGAGAGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCA 381
Db      315  CTTACGGGAGAGTGAAGAGAGAGAGAGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCA 374
Qy      382  TCAAGAGAGAGAGAGAGAGAGAGAGAGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTGA 441
Db      375  TCCGAGAGAGAGAGAGAGAGAGAGAGAGATCTGAGGAGTCTGAGGAGTCTGAGGAGTGA 434
Qy      442  TCAATGTTTCAATCAACACCCCAATCTATGCGATCTCAATGAGTCTTGAAGATGA 501
Db      435  TCAATGTTTCAATCAACACCCCAATCTATGCGATCTCAATGAGTCTTGAAGATGA 494
Qy      502  GCAAGATTTGATTTGTCATGAGAGTCTGAGGAGAGAGTCTGATGATTAATCAATCAAGT 561
Db      495  GCAAGATTTGATTTGTCATGAGAGTCTGAGGAGAGAGTCTGATGATTAATCAATCAAGG 554
Qy      562  AGCGGCAAGGCTGAGTGAAGAGAGAGAGAGAGTCTTCTGAGAGAGTCTGAGTCTGCGC 621
Db      555  AGCGGCAAGGCTGAGTGAAGAGAGAGAGAGAGTCTTCTGAGAGAGTCTGAGTCTGCGC 614
Qy      622  TGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 681
Db      615  TGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 674
Qy      682  TAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 741
Db      675  TAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 734
Qy      742  GCAAGTCTCTCAGAGAGTCTGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 801
Db      735  GCAAGTCTCTCAGAGAGTCTGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 794
Qy      802  GGAAGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 861
Db      795  GGAAGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 854
Qy      862  TGGTGCATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 921
Db      855  TGGTGCATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 914
Qy      922  GTAACGGGAGTTCACCTGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGTCTT 981
Db      915  GTAACGGGAGTTCACCTGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGTCTT 974
Qy      982  TGTTAATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 1041
Db      975  TGTTAATGAGAGAGAGAGAGAGAGAGAGAGTCTTCAAGTCTGAGAGAGAGAGAGTCTT 1034

```


QY 802 GGAAGCCCTATGTGAGGCGCCAGAGGTGAGACAGCTGCTCTGAGGCGTTCTCTGTACATCC 861
DB 856 GGAAGCCCTACACAGGCGCCAGAGGTGAGACAGCTGCTCTGAGGCGTTCTCTGTACATCC 915
QY 862 TGTGTGATGAGACCATGCTCTTGTGACGAGGAGATCATTAACATGTGTAAAGCAATCA 921
DB 916 TGTGTGATGAGACCATGCTCTTGTGACGAGGAGATCATTAACATGTGTAAAGCAATCA 975
QY 922 GTAAAGGAGGCTTACCGTGAAGCGCCCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCG 981
DB 976 GCAAGGAGGCTTACCGTGAAGCGCCCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCG 1035
QY 982 TGTAAATGTGAAACCCCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1041
DB 1036 TGTAAATGTGAAACCCCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1095
QY 1042 TCAACTGAGGTTACACCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1101
DB 1096 TCAACTGAGGTTACACCAAGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1155
QY 1102 CTGATGAGTGTGAGGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1161
DB 1156 CTGATGAGTGTGAGGCGGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1215
QY 1162 TCCTGAGAAATGAGCAAGGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1221
DB 1216 TCCTGAGAAATGAGCAAGGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1275
QY 1222 GCACTGATCTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1281
DB 1276 GCACTGATCTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1335
QY 1282 TGTGTCAAAATGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1341
DB 1336 TGTGTCAAAATGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1395
QY 1342 GCTTAAAGCTTCCGAAAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1401
DB 1396 ACCTCAAGCTTCCGAAAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1455
QY 1402 AGAGAGACCTTCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1461
DB 1456 AGAGAGACCTTCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1509
QY 1462 TATCTCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1521
DB 1510 TATCTCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1566
QY 1522 ACTCTCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1581
DB 1567 ACTCTCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1626
QY 1582 GTGAGGAGACCTTCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1641
DB 1627 GTGAGGAGACCTTCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1686
QY 1642 GCATTTCAAACTCAATGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1701
DB 1687 GCATTTCAAACTCAATGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1746
QY 1702 CTTTGTGCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1761
DB 1747 CTTTGTGCTGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1806
QY 1762 CAGGAGGCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1821
DB 1807 CAGGAGGCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1866
QY 1822 CTGAAGGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1881
DB 1867 CTGAAGGCTTCCAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1926

QY 1882 TTGAGAGGCTTCCCTCAAGAG-----GTCTGAAGCGATGTGTGAGGAGATCT 1929
DB 1927 TTGAGAGGCGTCCCTCAAGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1986
QY 1930 TGGGAGGAGGCTTCTTCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 1989
DB 1987 TGGGAGGAGGCTTCTTCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 2046
QY 1990 CCTTGAAGATCTGCTCAAGAGGCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 2049
DB 2047 CACTGAGGAGTGTCTCAAGAGGCTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 2105
QY 2050 GCTCTGAGAGGTTTGAAGAGAACTGCTGAGTGTCTCC 2092
DB 2106 GCTCTGAGATGACAGCTGTTGACCCGAGGAGGAGATGCTTC 2148

RESULT 12
US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids a
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803.10NCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 47.3%; Score 1372.2; DB 17; Length 2501;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGTGCGCCAGCCACTTCACCTGCGGCTCCGCGACCATGAGTGTGAGGCTTAC 141
DB 65 CTATTTGATTCCTCCGCGCCCTTGTCTACCTCTGCTGCGCATGAGTGTGAGGCTTAC 124
QY 142 TCCAGGCGCCGAGGAGGCTTCCGCGCTTCCGCGCTGCGCTGCGAGAGCGCCCGCGCG 201
DB 125 CGGCGCGCTCCGCGCCCACTTCCTGCGCGGAGAGCTA-----GCCCGCGCGC 172
QY 202 TGGCGGAGGAGCTCATGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 261
DB 173 TGGCGGAGGAGCTCATGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 232
QY 262 ACCATCAACACACACCTGCGGAGCGCTTACGAGTGTCTGAGAGCGTGTGAGGCGTGTGAGGCG 321
DB 233 ACCATCAACACACACCTGCGGAGCGCTTACGAGTGTCTGAGAGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 292
QY 322 CCTACGGAAGAGTGAAGAGGAGAGAGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 381
DB 293 CCTACGGAAGAGTGAAGAGGAGAGAGAGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 352
QY 382 TCAGGAAAGCAAAATCAAAAGTAGAGGAGTGTGTCATTCGACATTCGAGGAGGAGTTAGA 441

APPLICANT: KHAH, Farrah A.
 APPLICANT: RECIPON, Shirley A.
 APPLICANT: AZIMZAI, Yalda
 APPLICANT: POLICKY, Jennifer L.
 APPLICANT: DING, Li
 APPLICANT: GREYER, Megan
 APPLICANT: ELIOT, Vicki S.
 APPLICANT: THANGAVELU, Kavitha
 APPLICANT: BATRA, Sajeev
 APPLICANT: ISON, Craig H.
 TITLE OF INVENTION: HUMAN KINASES
 FILE REFERENCE: PI-0125 PCT
 CURRENT APPLICATION NUMBER: US/10/311,034
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
 60/228,056
 PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-07-25
 NUMBER OF SEQ ID NOS: 52
 SOFTWARE: PERL Program
 SEQ ID NO 38
 LENGTH: 3360
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CH1
 US-10-311-034-38

Query Match 47.2%; Score 1370.6; DB 17; Length 3360;
 Best Local Similarity 81.5%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 339; Indels 35; Gaps 4;

82 CTACTGTCGCGCCGACCACTGCTGCGGTCCTCCCGGACCATGAGTGGTGGCTTAC 141
 39 CTACTGATTCCTCCGCGCCCTTCTGCTCACTGCTGCGCATGAGTGGTGGTTCG 98
 142 TCCAGCGCCGCGGACGAGCTCTGCGGCTCGGCTCGGCTCGGAGAGCGCGCGCGC 201
 99 CGCGGCGCTCGGCGCCCACTCTCTCGGCG-----CAGAGCTAGCCCGCGCGC 145
 202 TGGCGGAGCGGCTCATGAGTGGCTTAACTCTGATGAGAGAGAGCGGCTGAGCGGC 261
 146 TGGCGGAGGCGCTATGAGTGGCGCAAGCGCTTAAAGAGAGAGCGGCTGAGAGCGGC 205
 262 ACCATCAAAACAACAACCTGCGGCAAGCTTGAAGTCTTGAAGAGCGTGGCGAAGGCA 321
 206 ACCACCAACAACAACCTGCGGCAAGCTTGAAGTCTTGAAGAGCGTGGCGAAGGCA 265
 322 CCTACGGGAGAGTGAAGAGGCAAGAGAGCTGGGCGCTGCTGGGCGATCAAGTCA 381
 266 CCTACGGGAGAGTGAAGAGGCGGAGAGCTGGGCGCTGCTGGGCGATCAAGTCA 325
 382 TCAGAGAAAGCAAAATCAAGATGAGAGTCTGCTGACATACGAGAGGAGATTGAGA 441
 326 TCCGAGAGGCAAAATCAAGATGAGAGTCTGCTGACATACGAGAGGAGATTGAGA 385
 442 TCATGCTTCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAAGATGCA 501
 386 TCATGCTTCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAAGATGCA 445
 502 GCAAGATTGATGTGATGATGAGATGAGAGAGCTGGGCGCTGCTGATTAATTAATCACTG 561
 446 GCAAGATTGATGTGATGAGATGAGAGAGCTGGGCGCTGCTGATTAATTAATCACTG 505
 562 AGCGGCAACGCTGATGAGAGGAGAGCGCAAGCAATTTCTTCCGACAGATCTGCTGCGC 621
 506 AGCGGCAACGCTGATGAGAGGAGAGCGCAAGCAATTTCTTCCGAGAGATCTGCTGCGC 565
 622 TGCATCTACTGCAACGAAGGAGATCTTCAACGAGATCTCAAGTGGAGAAACATCTTC 681
 566 TGCATCTACTGCAACGAAGGAGATCTTCAACGAGATCTCAAGTGGAGAAACATCTTC 625

682 TAGATGCCAATGGAACATCAAGATTGCTGACTTGGCTCTCTCCACCTGATCAACAAG 741
 626 TGGGTGCCAATGGAATATCAAGATTGCTGACTTGGCTCTCTCCACCTGATCAACAAG 685
 742 GCAAGTTCCTCCAGACCTTCTGAGGAGCCCTCTGAGGCTGAGTACGATCAAG 801
 686 GCAAGTTCCTCCAGACCTTCTGAGGAGCCCTCTGAGGCTGAGTACGATCAAG 745
 802 GGAAGCCCTTATGAGGAGCCAGAGGTGAGACAGTGTCTGAGGCTTCTCTGATCAATCC 861
 746 GGAAGCCCTTATGAGGAGCCAGAGGTGAGACAGTGTCTCTGAGGCTTCTCTGATCAATCC 805
 862 TGGTCATGACCAACATGCTTGAAGAGATGATCAATTAACATGAGTGAACAATCA 921
 806 TGGTCATGACCAACATGCTTGAAGAGATGATCAATTAACATGAGTGAACAATCA 865
 922 GTAAAGGAGCTTACCTGAGACCCGCGCAAGCTGCTGAGGCTGAGTGAAGGAGTGGC 981
 866 GCAAGGAGCTTACCTGAGACCCGCGCAAGCTGCTGAGGCTGAGTGAAGGAGTGGC 925
 982 TGTAAATGAGTGAACCCCAAGCTGAGGAGCACTGAGAGATGATCAATTAAGTGGG 1041
 926 TGTAAATGAGTGAACCCCAAGCTGAGGAGCACTGAGAGATGATCAATTAAGTGGG 985
 1042 TCAACTGAGGATTAACAACCGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101
 986 TCAACTGAGGATTAACAACCGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045
 1102 CTACTGATCTTGGCGGAGCTTCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161
 1046 CTGCACTGATCTTGGCGGAGCTTCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1105
 1162 TCCGAGAGATGAG 1221
 1106 TCCGAGAGATGAG 1165
 1222 GCACTGATCTTGGCGGAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1281
 1166 GCACTGATCTTGGCGGAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1225
 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTTCTTGGCGCTTGGCAAGAGCA 1341
 1226 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTTCTTGGCGCTTGGCAAGAGCA 1285
 1342 GCCTTAAGCTTCCGAG 1401
 1286 ACTCAAGCTTCCGAG 1345
 1346 AGAGAGAGCTTCCGAG 1399
 1462 TATCTGCTTCCGAG 1521
 1400 ---CCCTGCTTCCGAG 1456
 1522 ACTCTCTTCCGAG 1581
 1457 ACTCTCTTCCGAG 1516
 1582 GTGGGAG 1641
 1517 GTGGGAG 1576
 1642 GCATTTCAAACTCAATGAG 1701
 1577 GCATTTCAAACTCAATGAG 1636
 1702 CTTTGGCTTCCGAG 1761
 1637 CTTTGGCTTCCGAG 1696
 1762 CAGGGGCTGATGAG 1821

Db 1697 CAGGGGCTGTGAGGAGGAGCAATCCTGTCTCTGAGTCTCTTTGACCAAGCTGGAATTGGC 1756
Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGACAACCTTGAGGGGCGC 1881
Db 1757 CTGAACGGCTCCCAAGAGCCCACTGCGGGGCTGTGTCTGTGACAACCTTGAGGGGCGC 1816
Qy 1882 TTGAGCAGCCTCCCTCAAGAG-----GTCTGAAGCGATGTGTGGAGGAATCCT 1929
Db 1817 TTGAGAGAGCCCTCTCAAGAGGCGCTGGAAGCTGCTGAGGCGCTGGCGGAGATCCTT 1876
Qy 1930 TGGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGTACTGACCTTACAGACAG 1989
Db 1877 TGGGGGACAGCTGCTTTTCCCTGACAGACTGCCAAGAGTGTACTGACCTTACAGACAG 1936
Qy 1990 CCTTGAAGATCTGCTCAAGCTCAGCTGAGGAGGAGAGATGTCCTTATGATGAGGTAG 2049
Db 1937 CACTGAGGGTCTGCTCAAGCTCAGCTGAGTGTGAGTGTGACATGCCCCAG-CCGGGTACG 1995
Qy 2050 GCTCTGAGAGGGTTTGACAGAGAACCTGGGTGGATTCCTCC 2092
Db 1996 GCTCTGAGATGACAGCTGGTTGACCCCGAGGGAGATGCTTTC 2038

RESULT 14
US-09-963-159-3
; Sequence 3' Application US/09963159
; Patent No. US20020077312A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700. A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-5011
; CURRENT APPLICATION NUMBER: US/09/963,159
; PENDING FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-159-3

Query Match 46.9%; Score 1360.2; DB 9; Length 1884;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

Qy 169 CTTCCCGCCCTGGCTGGAGAGGCGCGCGCTGGGCGGACGGGCTCATCAAGTCGCTA 228
Db 35 CCACTCCCTCGGCGGAGAGCTAGCCGCGCTGGCGGAGGGCTGATCAAGTCGCCA 94
Qy 229 AACCTCTGATGAAGAACAGCGGCTGAGCGGACATCAACAACAACCTGGCGGACC 288
Db 95 AGCCCTTAAGAGAGACAGCGGCTGAGCGGACATCAACAACAACCTGGCGGACC 154
Qy 289 GCTACGAGTCTCTGAGAGCGCTGGGCAAGGCGACTTACGAGAGGTGAAGAGCGACAG 348
Db 155 GCTACGAGTCTCTGAGAGCGCTGGGCAAGGCGACTTACGAGAGGTGAAGAGCGCGGG 214
Qy 349 AGAGCTCGGGGCGCTGTGTGCGCATCAAGTCCATCAAGAAAGCAAAATCAAAATGAGC 408
Db 215 AGAGCTCGGGGCGCTGTGTGCGCATCAAGTCCATCAAGAAAGCAAAATCAAAATGAGC 274
Qy 409 AGGATCTGCTGCAATACGAGGAGGAGATTGATCATGTCTTCACTCAACCAACCCGACA 468
Db 275 AAGATCTGATGCAATACGAGGAGGAGATTGATCATGTCTTCACTCAACCAACCCGACA 334
Qy 469 TCATTGCTATCATGAAGTGTGTTGAAGTAAGACGAAGATTGATTGTATGATGATG 528
Db 335 TCATTGCTATCATGAAGTGTGTTGAAGTAAGACGAAGATTGATTGTATGATGATGATG 394

Qy 529 CCAAGCCGAGCGATCTGTATGATTATCATGAGCGGCGCAACCGGCTGATGAGGGGACG 588
Db 395 CCAAGCCGAGCGACCTTTATGATCAATACAGCGAGCGGCGACAGCTGATGAGGCGAAG 454
Qy 589 CCAAGCATTTCTTCCGACAGATGATGTCTGCGCTGACATCTGCGCACAGAACGGATGCG 648
Db 455 CTAGGATTTCTTCCGCAATGCTCTCTGCGGACATCTTGGCATTCAGAACAGATGG 514
Qy 649 TTCAACGAGATCTCAAGCTGGAAGAAATCTTCTGATGCCAATGAAATCAAGATTG 708
Db 515 TTCAACGAGATCTCAAGCTGGAAGAAATCTTCTGATGCCAATGAAATCAAGATTG 574
Qy 709 CTGACTTTGGCTCTCCAACTGTACACCAAGGCAAGTCTTCAACGCTTGTGAGGA 768
Db 575 CTGACTTTGGCTCTCCAACTGTACACCAAGGCAAGTCTTCTGAGCAATTCGTGGGA 634
Qy 769 GCCCTCTCAAGCGCTCGCTGAGATAGTCAACGGGAGGCGCTTATGTTGGGCGCAAGGTG 828
Db 635 GCCCTCTCAATGCTCGCGCAGAGATTTCTCATGGAAGCTTACACAGGCTCCAGAGGTG 694
Qy 829 ACAAGTGTCTTGGGCGTTCTCTGTACATCTTGTGATGACATGACATGCTTTGACG 888
Db 695 ACAAGTGTCTTGGGCGTTCTCTGTACATCTTGTGATGACATGACATGCTTTGATG 754
Qy 889 GGCAGGATCAATAAACAATGATGAAGAAATCAAGTCAAGGGGCTTACCTGAGCGGCCCA 948
Db 755 GGCAGGATCAATAAAGATCTGATGAAGAAATCAAGTCAAGGGGCTTACCTGAGCGGCCCA 814
Qy 949 AGCGTCCGATGCTGTGCGCTGATCCGATGCGGTGTTAATGTTGAAGCCCAACCGTGGG 1008
Db 815 AACCTCTGATGCTGTGCGCTGATCCGATGCGGTGTTAATGTTGAAGCCCAACCGTGGG 874
Qy 1009 CCAACTGAGAGATGTAAGCAATGATGTTGGGTCAACTGGGGTTTACCAACCGGAGTGC 1068
Db 875 CCAACTGAGAGATGTAAGCAATGATGTTGGGTCAACTGGGGTTTACCAACCGGAGTGC 934
Qy 1069 GGAAGCAAGAAAGCCCTGCGGAGGAGGAGGAGCACTGATGATGTTGGGCGGCGCTCCA 1128
Db 935 GAGAGCAAGAGGCTCGGCAATGAGGAGGAGGAGCACTGATGATGTTGGGCGGCGCTCCA 994
Qy 1129 TGGCGGATGTTACGTCGCTCTCGGCGCCCTCTCGAGAGAAATGAGCCAGGATGCA 1188
Db 995 TGGCTGATGTTGCTCGGCGTTCTCTCGGCGCCCTCTCGAGAGAAATGAGGCGGATGCA 1054
Qy 1189 GCTTCTTCAAGACAGAGTCCGAGAGGAGGAGCACTGATGATGTTGGGCGGCGGAC 1248
Db 1055 GCTTCTTCAAGACAGAGTCCGAGAGGAGGAGCACTGATGATGTTGGGCGGCGGAC 1114
Qy 1249 ATTCTCTTAAGAGTCCCGAAAGAGATGATGCTCAAAATCTGCAAGGTGACCCCG 1308
Db 1115 ATTCTCTTAAGAGTCCCGAAAGAGATGATGCTCAAAATCTGCAAGGTGACCCCG 1174
Qy 1309 CTGAGGATACCTCTTCTCGGCTCGGCAAGAGAGCTTAAAGTTCGAAAGGCAATTCTCA 1368
Db 1175 CTGATGACATGCTGCTGCTGCTGCGCAAGAGCACTCAAGCTGCGCAAGGCAATTCTCA 1234
Qy 1369 AGAAAGTCTCTTACCTCTGCTGAGGAGGATGACAGAGAGCCTCAAGAACTCAAGCCGG 1428
Db 1235 AGAAAGTCTCTTACCTCTGCTGAGGAGGATGACAGAGAGCCTCAAGAACTCAAGCCGG 1294
Qy 1429 TGCCTGATCTCAAGGAGAGCTCTGCTGCTGATCTCTGCTGCTCAAGGAGGATCC 1488
Db 1295 TGCCTGAGACCTCAAGGAGAGCTCTGCTGCTGATCTCTGCTGCTCAAGGAGGATCC 1345
Qy 1489 TTAAGAAGTCTGACAGCGGATGATGTTACTCTCTCAAGAGCCCAAGGAGTCTG 1548
Db 1346 TCAAGAGCCCGACAGCGGATGATGTTACTCTCTCAAGAGCCCAAGGAGTCTG 1405
Qy 1549 GGAAGCTTTTGAAGCGCAAGTATGTTTGTGAGTGGGAGCCCGTGAAGCAAGTCTC 1608
Db 1406 GGAAGCTTTTGAAGCGCAAGTATGTTTGTGAGTGGGAGTCCCAAGGAGCAAGGCTC 1465
Qy 1609 CACAGGCTTACAGGCTCTCTCTCAACCGCAAGGCAATTCTCAAACTCATGGAAGTCT 1668

Db 1466 CGCAGCTTCGAGGCTCTCTCTCCATGCGAAGGCGATCTCAAACTCAATGCGAAGTTCT 15255

Db 1669 CCGGACAGCCTTAGAAGGCACTACCCCTAGACACTTTTGGCTCCCTGAGCCAACTGGCT 1728

Db 1526 CCGAGACAGCCTTAGAGCTCGCGGCCCCCACCACCTTGGGCTCCTCGATGAATCTGGCC 15855

Db 1729 CCTCCGATCTGACAGCCCGGCCAGACCGCCCTCAAGGGGCTGTGAGTAGGACAGCAATCC 1788

Db 1586 CACCTCGCCCCCTGGCCCCGGGCGAGCGACCCCTCAAGGGGCTGTGAGCGAGGACAGCAATCC 16454

Db 1789 TGTCTCTCCGAGCTCTTTTGAACAATTTGACTTGGCTGAACGTCTTCCGAAACCCCATCTGA 1848

Db 1646 TGTCTCTGAATCTCTTTATGACCAAGCTTGACCTTGGCTGAACGGCTCCAGAGCCCCCATCTGC 17050

Db 1849 GGGGCTGTGTCTGTGTGACAACCTGAGAGGGGCTTTAGACAGCTTCCCTCAGAAAG----- 1902

Db 1706 GGGGCTGTGTGTCTGTGACAACCTCAAGGGGCTTTGAGAGAGCCCCCTCAGAGGGGCCCTG 17655

Db 1903 -----GTCATAAGGATGTGTGCGAGAAATCTTGGGGGATATAGCTGCTTTCTGTGACAG 1956

Db 1766 GAAAGCTGTGAGGGGCTGGCGGCGGAGAAATCTTTGGGGGAGACAGCTGCTTTCCCTGAAG 18255

Db 1957 ACTGCGAAGAGGTGACTGACGCTTACAGACAAGCCCTTAGAATCTGCTCAAGCTCAAGC 2015

Db 1826 ACTGCGAAGAGGTGACAGCACTTACCGACAGGCACTAAGGCTCTGCTCAAGCTCAAGC 1884

RESULT 15
US-10-423-543-45
Sequence 45, Application US/10/423543
Publication No. US20040058355A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Libermann, Rosana K.
APPLICANT: Hunter, John J.
APPLICANT: Meyers, Rachel E.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Olandt, Peter J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Galvin, Katherine M.
APPLICANT: Chun, Miyoung
APPLICANT: Williamson, Mark J.
APPLICANT: Silos-Santiago, Immaculada
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
TITLE OF INVENTION: 18610, 33217, 21967, 11983, 11983, 38555 OR 593 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MP103-0230NM1M
CURRENT APPLICATION NUMBER: US/10/423,543
CURRENT FILING DATE: 2003-04-25
PRIOR APPLICATION NUMBER: US 10/278,036
PRIOR FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: US 09/711,216
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/205,447
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 10/012,055
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/248,325
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 10/003,690
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/248,893
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/797,039
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186,061
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 10/217,168
PRIOR FILING DATE: 2002-08-12

```

; Remaining Prior Application data removed - See File Wrapper or PALM
;
; NUMBER OF SEQ ID NOS: 119
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 45
;
; LENGTH: 1884
;
; TYPE: DNA
;
; ORGANISM: Homo Sapiens
;
US-10-423-543-45

```

Query Match	46.9%;	Score 1360.2;	DB 17;	Length 1884;
Best Local Similarity	84.2%;	Pred. No. 0;		
Matches 155;	Conservative	0;	Mismatches 273;	Indels 21;
				Gaps 2;

OY	169	CCTCCGCGCTG3CCTCGCAGAGAGCCCGCGCGCTGCGCGAGCGGGCTCATCAAGTCGCGCTA	228
Db	35	CCAATCCCTCGCGCCGCAAGAGCTAGCCCGCGCGCTGCGCGAGAGGGCTCATCAAGTCGCGCCA	94
OY	229	AACCTTGATGAGAGACGAGCGGCTGAGCGCGCACCATCAACAAACACACTCTGCGGCAAC	288
Db	95	AGCCCTTAATGAGAGAGACGAGCGGCTGAGAGCGGCACACACAAAGCACACACTGCGGCAAC	154
OY	289	GCTACGAGTTCTCTGAGACGCTGGGCGCAAGGGCACCTACGGGAAAGGTAAAGAGCAAGAG	348
Db	155	GCTACGAGTTCTCTGAGACCTCTGGGCAAAAGGCACCTACGGGAAAGGTAAAGAGAGCGCGGG	214
OY	349	AGAGCTCGGGGCGCTCTGCTGGCCATCAAGTCCATCAGAGAAAGACAAATCAAAAGATAGC	408
Db	215	AGAGCTCGGGGCGCGCTGCTGGCCATCAAGTCAATCCGGAAGAGCAAAATCAAGATGAGC	274
OY	409	AGAGTCTGCTGCACATACGAGAGGAGATTGAGATCAATGTCTTCACTCAACACCCCCACA	468
Db	275	AAGATCTGATGACATACGAGAGGAGATTGAGATCAATGATCATCACTCAACCACTTCA	334
OY	469	TCATTGCAATCCATGAGTAGTGTGAGAAATAGCAGCAAGATTGATGATCATGAGATGATG	528
Db	335	TCATTGCAATCCATGAGTAGTGTGAGAAACAGCAGCAAGATCGTGAATGTCATGAGAGATG	394
OY	529	CCAGCCGAGCGGATCTGTATGATTAATCATCATGAGCGGCGCAACGCTGATGATGCGCGGACG	588
Db	395	CCAGCCGAGGCGGACCTTTATGACTATCACTACGAGCGGCGGACAGCTCATGATGAGCGCAAG	454
OY	589	CCAGGCAATTTCTTCCGACAGATCGTGTGCTGCGCCGCGACCTTCTGCGCACAGAGCGGATCG	648
Db	455	CTAGGCAATTTCTTCCGGAAGATCGTCTTGTGCGGTGACCTTATGCGCATCAGAACAGATTGG	514
OY	649	TTCAACGAGATCTTCAAGCTGGAAAAACATCTCTTGAATGCGAATGAGAAACATCAAGATTG	708
Db	515	TTCAACCGAGATCTTCAAGCTGGAGAAACATCTCTTGAATGCGAATGAGAAATATCAAGATTG	574
OY	709	CTGACTTTGGGCTCTCCAACTGTGACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGTGGGA	768
Db	575	CTGACTTGGGCTCTCCAACTTACCATCAAGGCAAGTTCTGTGACAAACATTCGTGTGGGA	634
OY	769	GCCTCTCTACGCTCGCTCGATGATATCAACGGGAGACCTTATGTGTGGCCACAGAGCTGG	828
Db	635	GCCTCTCTACGCTCTCGCAGAGATTTCTCATATGGGAGCCCTTACACAGGCCACAGAGGTTGG	694
OY	829	ACAGCTG3GTCTGTGGGCGTTCTCTGTACATCTGTGTGATG3CATG3CATCCATCCCTTTGACG	888
Db	695	ACAGCTG3GTCTGTGGGCGTTCTCTCTTACATCTGTGTGATG3CATG3CATCCATCCCTTTGATG	754
OY	889	GGCAGGATCATTAACACTGGTGAAGAAATCATGTAACGGGGGCTTACCGTGAAGCGGCCCA	948
Db	755	GGCAGTGAACATTAACATCTTATGTAAACAGATCAGCAACGGGGGCTTACCGG3AGCCACTTA	814
OY	949	AGCGGTCCGATGCTGTGGCGCTGATCCGGTGGCTGTTAATGTGTGAACCCCAACCGGTGGG	1008
Db	815	AACCTCTGATGCTGTGGCGCTGATCCGGTGGCTGTTGATGTGTGAACCCCAACCGCGGG	874
OY	1009	CCACAATGGAGATGTAACGAGTCAATTGGTGGGTCAACTGGGGGTTAACACACGG3AATGCG	1068
Db	875	CCACCCTGAAGATGTGCGCAGTACTGCTGTGGGTCAACTGGGGGTTACGCCACCCG3AATGCG	934

QY 1069 GGGAAAGAGAGAGCCCTGGTGAAGGGTGGGCAACCTAGTGTGACTTTGGCCGGGCTTCCA 1128
DB 935 GAGAGCAGAGAGCTCCGATGAGGGTGGGCACTTGGCATCTTGGCCGCTTCCA 994
QY 1129 TGGCGGACTGGTTACGTGCTCTCGGCCCCCTCTCTGAGATGATGAGCCAAAGGTGTGCA 1188
DB 995 TGGCTGACTGGCTCCGGGTTCTCTCCGCCCTCTCTGAGATGATGGGCGCAAGGTGTGCA 1054
QY 1189 GCTTCTTCAAGCAGCAGTGCCTGGAGGTGAGACTGTACTTGGGCTGGAGCGGCAAC 1248
DB 1055 GCTTCTTCAAGCAGCAGTGCCTGGAGGTGAGACTGTACTTGGGCTGGAGCGGCAAC 1114
QY 1249 ATTCTTTAAGAGTCCCGAAAGGAGATGACATGGCTCAAAATCTGCAAGTGAACCCGG 1308
DB 1115 ATTGCTCAAGAGTCCCGAAAGGAGATGACATGGCTCAAAATCTGCAAGTGAACCCGG 1174
QY 1309 CTGAGATTAACCTTCTTCTGCTCTGCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATTCTCA 1368
DB 1175 CTGATGACACTGCCCCATCGCCCTGGCAAGAGCACTCAAGCTGCCAAAGGCAATTCTCA 1234
QY 1369 AGAAAAAGTCTTACTCTCTGAGGGAGGTACAGAGAGACCTCAGGAATCAGACCGG 1428
DB 1235 AGAAGAGGTGTGAGCTCTGCAAGAGGGGTACAGAGAGACCTCAGGAGCTCAGCCAA 1294
QY 1429 TGCTGATTAATCCAGGGAGGCTGTCTCTGTATCCCTGCTGCCAAAGGAAAGGCAATCC 1488
DB 1295 TCCCTGGAGCCAGGGGAGGCTGTCTCTCTGTATCCCTGCTGCCAAAGGAAAGGCAATCC 1345
QY 1489 TTAAAGATCTGCAAGCAGTGAATCTGATTAATCTCTCTGCAAGCAGCAGTGTG 1548
DB 1346 TCAAGAGAGCCCGAGAGGAGGTCTGGCTACTACTCTCTCCGAGCCCGAGTGAATCTG 1405
QY 1549 GGGAACTTTAAGAGCAGTGAATGTGTTGTGAGTGGGAGCCCGTGAAGCAAGTCTC 1608
DB 1406 GGGAGCTTTGAGAGCAGGAGCAGTGTGTTGTGAGTGGGAGTCCCAAGAGCAGAACTTC 1465
QY 1609 CACAGGCTTCAAGGCTCTCTCTTCCACGCAAGGCAATTCCTCAATGCGCAAGTCTC 1668
DB 1466 CGAAGCTTCAAGGCTCTCTCTTCCATGCAAGGCAATTCCTCAATGCGCAAGTCTC 1525
QY 1669 CCCGACAGCCTTAAGAGCACTACCCCTAGCACTTTGGCTCCCTGGAACCACTGGGCT 1728
DB 1526 CCCAGACAGCCTTAAGAGCTCGCGGCCCCCACTTCCGCTCCCTGATGAACTCGGCC 1585
QY 1729 CTTCCATCTGTCAGCCCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAAGACAGATCC 1788
DB 1586 CACCTCGCCCCCTGGCCCGGCGCAGCCGACCTCAGGGGCTGTGAGGAGACAGATCC 1645
QY 1789 TGTCTCCGAGTCTTGAACCAATTGACCTTGCCTGAACGTCCTCCGAAACCCCACTGA 1848
DB 1646 TGTCTCTGAGTCTTGAACCAAGTGAACCTTGTGAAAGGCTCCAGAGCCCCCACTGC 1705
QY 1849 GGGGCTGTGTGTGTGTGAGCAACTGAGGGGGCTTGAAGAGCCTCCCTCAGAAAG----- 1902
DB 1706 GGGGCTGTGTGTGTGTGAGCAACTGAGGGGGCTTGAAGAGCCTCCCTCAGAGGGCCCTG 1765
QY 1903 -----GTGAAAGCGATGTGTGAGGAATCTTGGGGGATAGCTGCTTTTCTGTACAG 1956
DB 1766 GAAAGTGTGTGTGTGTGAGGGCTGTGAGGAGATCTTTGGGGGACAGCTGCTTTTCCGTACAG 1825
QY 1957 ACTGCCAAGAGGTGACTGAGCAGCCTTACAGACAGAGCCCTAGGAATGTGCTCAAGCTCAGC 2015
DB 1826 ACTGCCAAGAGGTGACTGAGCAGCCTTACAGACAGGCACTGAGGGTGTGCTCAAGCTCAGC 1884

Search completed: May 16, 2005, 09:29:15
Job time : 1661.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 11:55:01 ; Search time 464.804 Seconds
(without alignments)
10216.077 Million cell updates/sec

Title: US-09-980-464-4
Perfect score: 2902
Sequence: 1 cactagtgcgtccaagat.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 beqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PRTUS.COMB.seq.*
6: /cgn2_6/prodata/1/ina/bocktile1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	US-09-579-664B-4	Sequence 4, Appl1
2	2902	100.0	2902	US-10-355-975A-4	Sequence 4, Appl1
3	1372.2	47.3	2501	US-09-799-451-115	Sequence 215, Appl1
4	204.2	7.0	2224	US-09-949-016-3384	Sequence 2384, Appl1
5	204.2	7.0	2224	US-09-949-016-2385	Sequence 2385, Appl1
6	204.2	7.0	2950	US-09-949-016-1546	Sequence 1546, Appl1
7	204.2	7.0	2950	US-09-949-016-1547	Sequence 1547, Appl1
8	202.6	7.0	2175	US-09-984-890-1	Sequence 1, Appl1
9	202.6	7.0	2175	US-10-274-194-1	Sequence 1, Appl1
C 10	189.2	6.5	484	US-09-270-767-1257	Sequence 1257, Appl1
C 11	189.2	6.5	484	US-09-270-767-16539	Sequence 16539, Appl1
12	186.4	6.4	2908	US-09-930-181-1	Sequence 1, Appl1
13	182.4	6.3	2652	US-08-557-006C-3	Sequence 29, Appl1
14	182.4	6.3	2761	US-08-557-006C-24	Sequence 24, Appl1
15	181.2	6.2	1742	US-08-557-006C-38	Sequence 38, Appl1
16	179.8	6.2	2698	US-08-677-298-1	Sequence 1, Appl1
17	179.8	6.2	2912	US-09-949-016-3363	Sequence 2363, Appl1
18	178.2	6.1	2914	US-09-949-016-343	Sequence 143, Appl1
19	176.4	6.1	2899	US-09-774-528-112	Sequence 112, Appl1
20	174	6.0	3364	US-09-930-181-3	Sequence 3, Appl1
21	173	6.0	2334	US-10-003-680-3	Sequence 3, Appl1
22	173	6.0	2337	US-10-116-326-1	Sequence 1, Appl1
23	172	6.0	2980	US-10-003-690-1	Sequence 1, Appl1
24	172.2	5.9	1539	US-09-633-328B-1	Sequence 1, Appl1
25	172	5.9	3609	US-09-799-875-6	Sequence 6, Appl1
26	172	5.9	5983	US-09-799-875-4	Sequence 4, Appl1
27	171.4	5.9	1647	US-09-101-146-44	Sequence 44, Appl1

28	165.6	5.7	1929	3	US-09-359-161.4	Sequence 44, April
29	153.6	5.6	1747	3	US-08-557-006C-44	Sequence 44, April
29	153.6	5.6	1747	3	US-08-557-006C-44	Sequence 44, April
30	161.6	5.6	1663	4	US-09-949-016-3303	Sequence 4303, April
31	161.6	5.5	1507	3	US-09-522-889.4	Sequence 4, April
32	160.4	5.5	1544	3	US-09-734-673.1	Sequence 1, April
33	160.4	5.5	2060	3	US-09-522-849.1	Sequence 1, April
33	158	5.4	1766	3	US-08-557-006C-37	Sequence 37, April
34	158	5.4	1766	3	US-08-557-006C-37	Sequence 37, April
35	158	5.4	1783	3	US-08-557-006C-36	Sequence 36, April
36	155.6	5.4	1051	3	US-09-522-849.3	Sequence 3, April
37	154	5.3	2112	4	US-10-115-136.5	Sequence 5, April
38	149.4	5.1	1599	3	US-09-256-465.1	Sequence 1, April
39	149.4	5.1	1599	3	US-09-167-322.3	Sequence 3, April
40	149.4	5.1	1599	4	US-09-023-655.5-1004	Sequence 1004, April
41	147.8	5.1	1254	4	US-09-590-740.3	Sequence 3, April
42	147.8	5.1	1281	4	US-09-417-197-70	Sequence 70, April
43	144.4	5.0	2184	4	US-09-471-197-138	Sequence 138, April
44	144.4	5.0	2610	2	US-09-421-771.1	Sequence 1, April
45	144.4	5.0	2610	3	US-09-091-058-1	Sequence 1, April

ALIGNMENTS

RESULT 1
 US-09-579-664B-4
 : Sequence 4, Application US/09579664B
 : Patent No. 6514719
 : GENERAL INFORMATION:
 : APPLICANT: Immunex Corporation
 : APPLICANT: Bird, Timothy A.
 : APPLICANT: Virca, G. Duke
 : APPLICANT: Martin, Unja
 : APPLICANT: Anderson, Dick M.
 : TITLE OR INVENTION: NOVEL MORINE AND HUMAN
 : FILE REFERENCE: 2223-A
 : CURRENT APPLICATION NUMBER: US/09/579,664B
 : CURRENT FILING DATE: 2000-05-26
 : NUMBER OF SEQ ID NOS: 36
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 4
 : LENGTH: 2902
 : TYPE: DNA
 : ORGANISM: Mus musculus
 : US-09-579-664B-4

Query Match	100.0%;	Score 2902;	DB 4;	Length 2902;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 2902;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	CACTAGTGATCCAAAGATTGCGCAGAGCGTGTGGGTGCGGCTGTGACCTTGAAG	60	
Db	1	CACTAGTGGATCCAAAGATTGCGCAGAGCGTGTGGGTGCGGCTGTGACCTTGAAG	60	
QY	61	CCCGCGGCTCAGCGCGCGCTGTCTACTGTGTGCCCGAACCATCTCACTTGGCGTCCCCGCA	120	
Db	61	CCCCCGGCTCAGCGCGCGCTGTCTACTGTGTGCCCGAACCATCTCACTTGGCGTCCCCGCA	120	
QY	121	CCATGGAGTGGGTGAGCTTACTCCAGGCGCCGAGCGCAGGCTCCCTCGGCTCCCGCCCTGG	180	
Db	121	CCATGGAGTGGGTGAGCTTACTCCAGGCGCCGAGCGCAGGCTCCCTCGGCTCCCGCCCTGG	180	
QY	181	CCTCGGAGAGCGCCCGCGCGCTGTGCGGAGCGGGCTCATCAATTCGCTTAACTCTGTATGA	240	
Db	181	CCTCGGAGAGCGCCCGCGCGCTGTGCGGAGCGGGCTCATCAATTCGCTTAACTCTGTATGA	240	
QY	241	AGAGCGAGCGGTGAAAGCGGCAACATCAAAACACACCTTGGCGCACCGCTTACGATTC	300	
Db	241	AGAGCGAGCGGTGAAAGCGGCAACATCAAAACACACCTTGGCGCACCGCTTACGATTC	300	
QY	301	TGGAGACGCTGGGCAAGGCACTTACGAGGAAGTGAAGAAAGGCAAGAGAGCTTGGGGC	360	
Db	301	TGGAGACGCTGGGCAAGGCACTTACGAGGAAGTGAAGAAAGGCAAGAGAGCTTGGGGC	360	

Db	2551	CTGAGTACCCAGATGATGATGATGCTCTGTTTCTGAGGCCAACGGACCTAGAAATGTC	2580
Qy	2551	TGACTTATTTATTTTTTTGTGATTTCTCACTCTGTGTTTTTGGTTTTGTGTTTGTG	2640
Db	2561	TGACTTATTTATTTTTTTGTGATTTCTCACTCTGTGTTTTTGGTTTTGTGTTTGTG	2640
Qy	2641	TTTTTGTTTTAAAGTAATTTTGTCTGCTTTCAATAATGTAATGCTGTGTTCTGGGAAAC	2700
Db	2641	TTTTTGTTTTAAAGTAATTTTGTCTGCTTTCAATAATGTAATGCTGTGTTCTGGGAAAC	2700
Qy	2701	TCACATGTGCACTGAAAGTTATATGACAGAGAAGATTTGGCAATGATGTCCTCTATTC	2760
Db	2701	TCACATGTGCACTGAAAGTTATATGACAGAGAAGATTTGGCAATGATGTCCTCTATTC	2760
Qy	2761	AAGGGGGGGTGGGGGGGCTTTTTCAAATGATATGCTTTGAGACATGTCTGATTTGAGTCTCA	2820
Db	2761	AAGGGGGGGTGGGGGGGCTTTTTCAAATGATATGCTTTGAGACATGTCTGATTTGAGTCTCA	2820
Qy	2821	GTCCCTTACACCCCAAGGCTGGCCACCTCCCTCATCTTCATCTGTGGCCAAAAAAA	2880
Db	2821	GTCCCTTACACCCCAAGGCTGGCCACCTCCCTCATCTTCATCTGTGGCCAAAAAAA	2880
Qy	2881	AAAAAAAAAAAAAAAAAAAAAAAA 2902	
Db	2881	AAAAAAAAAAAAAAAAAAAAAAAA 2902	

```

RESULT 2
US-10-355-975A-4
; Sequence 4, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Vircsa, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CAMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355,975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-355-975A-4

```

Query Match	100.0%	Score 2302;	DB 4;	Length 2902;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2302; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	CAC	AGT	GAT	GC	AAA	GAA	TTC	GCG	CA	GAG	GCG	GTG	CTG	GCG	CTG	GAT	CA	CT	TAG	60
Db	1	CAC	AGT	GAT	GC	AAA	GAA	TTC	GCG	CA	GAG	GCG	GTG	CTG	GCG	CTG	GAT	CA	CT	TAG	60
Qy	61	CCC	CGG	GTC	CA	GCG	GCG	GCT	CTA	CTG	TGC	CCG	CCA	CCA	CTC	CA	CT	CGG	GCT	CCCC	120
Db	61	CCC	CGG	GTC	CA	GCG	GCG	GCT	CTA	CTG	TGC	CCG	CCA	CCA	CTC	CA	CT	CGG	GCT	CCCC	120
Qy	121	CCAT	GAG	TG	CGT	GGC	CTT	TAC	TCC	AC	CGC	CCG	CA	GCG	CA	GCG	CA	GCG	CA	CTT	180
Db	121	CCAT	GAG	TG	CGT	GGC	CTT	TAC	TCC	AC	CGC	CCG	CA	GCG	CA	GCG	CA	GCG	CA	CTT	180
Qy	181	CCT	GGA	GAG	GCG	CCG	GCG	CGT	GCG	GCG	CGA	CGG	GCT	CAT	CA	AGT	CGC	CTT	AA	240	
Db	181	CCT	GGA	GAG	GCG	CCG	GCG	CGT	GCG	CGA	CGG	GCT	CAT	CA	AGT	CGC	CTT	AA	240		
Qy	241	AGG	AC	GAG	CGG	TCG	GAC	CGC	CA	CTT	CA	CT	CA	CT	CA	CT	CA	CT	CA	300	
Db	241	AGG	AC	GAG	CGG	TCG	GAC	CGC	CA	CTT	CA	CT	CA	CT	CA	CT	CA	CT	CA	300	

QY	301	TGAGACGCTGGGCAAGGGCAAGCTTACCGGAAAGGTGAAGAAAGCACGAGAGAGCTCGGGG	360
Db	301	TGAGACGCTGGGCAAGGGCAAGCTTACCGGAAAGGTGAAGAAAGCACGAGAGAGCTCGGGG	360
QY	361	GTCTGTGGCCATCAATGTCATCAGGAAAGCAAAATCAAAGATGACAGAGATCTGCTGC	420
Db	361	GTCTGTGGCCATCAATGTCATCAGGAAAGCAAAATCAAAGATGACAGAGATCTGCTGC	420
QY	421	ACATACGAGGAGAGATTGAGATCATGTCTTCACTCAACCACTCCCATCATCTTGGCATCC	480
Db	421	ACATACGAGGAGAGATTGAGATCATGTCTTCACTCAACCACTCCCATCATCTTGGCATCC	480
QY	481	ATGAAGGTGTGAGAAATGACGCAAGATTGTGATTGACAGAGATGACGAGCGAGAGG	540
Db	481	ATGAAGGTGTGAGAAATGACGCAAGATTGTGATTGACAGAGATGACGAGCGAGAGG	540
QY	541	ATCTGTATGATTAATCATCATGAGCGGCGCACGCGTGAATGAGCGGAGCGCCAGGATTTCT	600
Db	541	ATCTGTATGATTAATCATCATGAGCGGCGCACGCGTGAATGAGCGGAGCGCCAGGATTTCT	600
QY	601	TCCGACAGATCGTGTCTGCGCTTGCACTACTGCGCACCAAGACGGGATGTCTTACCGAGATC	660
Db	601	TCCGACAGATCGTGTCTGCGCTTGCACTACTGCGCACCAAGACGGGATGTCTTACCGAGATC	660
QY	661	TCAAGCGTGAAGAAACATCCTCTTGAATGSCAAATGGAACATCAAGATTGCTGACCTTGGCC	720
Db	661	TCAAGCGTGAAGAAACATCCTCTTGAATGSCAAATGGAACATCAAGATTGCTGACCTTGGCC	720
QY	721	TCTCAACCTGTACCAAAAGGCAAGTTCCTCCAGAGCTTCTGAGGAGCCCTCTTACG	780
Db	721	TCTCAACCTGTACCAAAAGGCAAGTTCCTCCAGAGCTTCTGAGGAGCCCTCTTACG	780
QY	781	CCTGCGCTGAGATAGTCAACGGGAAAGCCATATGTGGGCCCAAGAGTGAACAGCTGTGCTC	840
Db	781	CCTGCGCTGAGATAGTCAACGGGAAAGCCATATGTGGGCCCAAGAGTGAACAGCTGTGCTC	840
QY	841	TGGGCGTTCCTCTGTATCATCTCTGTGTGATGAGCACATGACCTTTGAGCGGCAAGATATA	900
Db	841	TGGGCGTTCCTCTGTATCATCTCTGTGTGATGAGCACATGACCTTTGAGCGGCAAGATATA	900
QY	901	AAACACTGAGTGAAGCAAAATGATGTAACGGGGCTTACCGTGAAGCGGCCCAAGCCGTCGATG	960
Db	901	AAACACTGAGTGAAGCAAAATGATGTAACGGGGCTTACCGTGAAGCGGCCCAAGCCGTCGATG	960
QY	961	CCTGTGGCTGATCCGGTGGCTGTAAATGTGTGAACCCACACCGTCTGGGCAACACTGAGG	1020
Db	961	CCTGTGGCTGATCCGGTGGCTGTAAATGTGTGAACCCACACCGTCTGGGCAACACTGAGG	1020
QY	1021	ATGTAGCCATGATTTGTGTGGGTCAACTGGGGTTTACACACCGGAGTGTGGGAAACAGGAG	1080
Db	1021	ATGTAGCCATGATTTGTGTGGGTCAACTGGGGTTTACACACCGGAGTGTGGGAAACAGGAG	1080
QY	1081	CCCGGCGTGAAGGAGTGGGCAACCTAGTGTGTATCTTTGGCGGGGCTTCATGCGGACTGGT	1140
Db	1081	CCCGGCGTGAAGGAGTGGGCAACCTAGTGTGTATCTTTGGCGGGGCTTCATGCGGACTGGT	1140
QY	1141	TACGTGCGCTCTCGCGGCCCTCTCTGTGAGATGAGCAAGTGTGTGACCTTTCAAGC	1200
Db	1141	TACGTGCGCTCTCGCGGCCCTCTCTGTGAGATGAGCAAGTGTGTGACCTTTCAAGC	1200
QY	1201	AGCAGGTGCGGAGAGTGAAGCACTGTACTCTGGGCTGGAAGCGGCAACATTTCTTTAAGA	1260
Db	1201	AGCAGGTGCGGAGAGTGAAGCACTGTACTCTGGGCTGGAAGCGGCAACATTTCTTTAAGA	1260
QY	1261	AGTCCCGGAAAGGAGAAATGAATGCTCTCAAAATCTGCAAGGATGACCGGCTGAGAGATACCT	1320
Db	1261	AGTCCCGGAAAGGAGAAATGAATGCTCTCAAAATCTGCAAGGATGACCGGCTGAGAGATACCT	1320
QY	1321	CTTCTCGCCCTGTGCAAGAGCAGCTTTAAGCTTCCGAAGGCAATTTCAAGAAAAAGTCT	1380
Db	1321	CTTCTCGCCCTGTGCAAGAGCAGCTTTAAGCTTCCGAAGGCAATTTCAAGAAAAAGTCT	1380
QY	1381	CTACTCTGTACAGGAGAGGTACAGAGAGACCTTCAAGACTCAAGACCGGTCTGTATATCTC	1440

Db 1381 CTACCTGTGACGGGAGGTACAGAGAGACCTTCAAGAACTAGACCGGTGCTGATCTC 1440
Qy 1441 CAGGCGAGCCTGTCCCTGTGTATCCCTGTCTCCCAAGAAAAGCATCTTAAAGATCTC 1500
Db 1441 CAGGCGAGCCTGTCCCTGTGTATCCCTGTCTCCCAAGAAAAGCATCTTAAAGATCTC 1500
Qy 1501 GACGCGGTGATCTGTGTACTACTCTCTCCAGAGCCCAAGAGATCTGGGAACTCTTAA 1560
Db 1501 GACGCGGTGATCTGTGTACTACTCTCTCCAGAGCCCAAGAGATCTGGGAACTCTTAA 1560
Qy 1561 ACCGCACTGATGTGTGTGTAGTGGGAGCCCGTGTAGCAGAGATCTCAAGGCTTCAAG 1620
Db 1561 ACCGCACTGATGTGTGTGTAGTGGGAGCCCGTGTAGCAGAGATCTCTCAAGGCTTCAAG 1620
Qy 1621 GGTCTCTCTTCACCGAGAGGAGCATTTCAAACTCAATGGCAAGTTCTCCGCAAGCT 1680
Db 1621 GGTCTCTCTTCACCGAGAGGAGCATTTCAAACTCAATGGCAAGTTCTCCGCAAGCT 1680
Qy 1681 TAGAAGGCACTACCCCTTAGACCTTTGGCTCCCTGTAGACCACTGGCTCTCTCCATCTG 1740
Db 1681 TAGAAGGCACTACCCCTTAGACCTTTGGCTCCCTGTAGACCACTGGCTCTCTCCATCTG 1740
Qy 1741 CAGCCCGGCGCCAGCGCCCTCAGGGGCTGTAGTGAAGAGCAGGATCTGTCTCCAGT 1800
Db 1741 CAGCCCGGCGCCAGCGCCCTCAGGGGCTGTAGTGAAGAGCAGGATCTGTCTCCAGT 1800
Qy 1801 CCTTGAACCAATTGGAATCTGTGCTGAAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
Db 1801 CCTTGAACCAATTGGAATCTGTGCTGAAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
Qy 1861 CTGTGGAACAACCTGAGAGGGGCTTGAAGAGCTCCCTCAGAAAGGTCTGAAGGATGTGTGC 1920
Db 1861 CTGTGGAACAACCTGAGAGGGGCTTGAAGAGCTCCCTCAGAAAGGTCTGAAGGATGTGTGC 1920
Qy 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGTGCAAGAGGTGACTGAGCT 1980
Db 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGTGCAAGAGGTGACTGAGCT 1980
Qy 1981 ACAGACAAGCCCTTAGAATCTGCTCAAAAGCTCAGCTGAGAGAGGAGATGTGTGCTTAGT 2040
Db 1981 ACAGACAAGCCCTTAGAATCTGCTCAAAAGCTCAGCTGAGAGAGGAGATGTGTGCTTAGT 2040
Qy 2041 ATGGGGTGAAGCTCTGAGAGGGTCTTGGAGAGAACCTGGGTGGATTCCTCAAGTAATA 2100
Db 2041 ATGGGGTGAAGCTCTGAGAGGGTCTTGGAGAGAACCTGGGTGGATTCCTCAAGTAATA 2100
Qy 2101 GAGTACATCAAGGGCTCTAGCTCTGACGCTGACCTGAACCTGAAGATGAGAGAAATGCG 2160
Db 2101 GAGTACATCAAGGGCTCTAGCTCTGACGCTGACCTGAACCTGAAGATGAGAGAAATGCG 2160
Qy 2161 ATTGATGTGAGAAAGAAATGGAGACCTTGTGCTGCCGAGTGTATATGTGGGTGCTGAA 2220
Db 2161 ATTGATGTGAGAAAGAAATGGAGACCTTGTGCTGCCGAGTGTATATGTGGGTGCTGAA 2220
Qy 2221 GGTGCTTACTCTCTTGTGCTGCAATAGTGTCAACCATGATTCCTCCACCTGTGTCTGG 2280
Db 2221 GGTGCTTACTCTCTTGTGCTGCAATAGTGTCAACCATGATTCCTCCACCTGTGTCTGG 2280
Qy 2281 CTGCACTTCAATAGATTTTCTGTTTCATCAACCAAGGGTTAGAACCTGACTTCT 2340
Db 2281 CTGCACTTCAATAGATTTTCTGTTTCATCAACCAAGGGTTAGAACCTGACTTCT 2340
Qy 2341 GGGAGGTATGTGTAGTACTGCTGCAATTTAGAGAGAGAAACAGCTGTGTTTCAATCT 2400
Db 2341 GGGAGGTATGTGTAGTACTGCTGCAATTTAGAGAGAGAAACAGCTGTGTTTCAATCT 2400
Qy 2401 CTGTGCTGTGATCTCAAAAGACTGTGAGAGACTGTGAGCCGTGTGACTTCAATCTCAA 2460
Db 2401 CTGTGCTGTGATCTCAAAAGACTGTGAGAGACTGTGAGCCGTGTGACTTCAATCTCAA 2460
Qy 2461 GGGGACAGATGCCCTGAGACCCGATTTAGATCTCAGAGCTTGAAGCTGT 2520

Db 2461 GGGGACAGATGCCCTGAGACCCGATTTAGATCTCAGAGACTTGAACCTTGAAGCTGT 2520
Qy 2521 CTAAGTACCAAGATGTGATGTAGTCTGTGTTTCTAGGCGCAACGGGACCTAGATGTGC 2580
Db 2521 CTAAGTACCAAGATGTGATGTAGTCTGTGTTTCTAGGCGCAACGGGACCTAGATGTGC 2580
Qy 2581 TGAATTAATTTTGTGATTTCTCACTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2640
Db 2581 TGAATTAATTTTGTGATTTCTCACTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2640
Qy 2641 TTTTGTGTTTAAAGTAAATTTGCTGTTCAATATGTAATGCTGTGTTTGGGAGAC 2700
Db 2641 TTTTGTGTTTAAAGTAAATTTGCTGTTCAATATGTAATGCTGTGTTTGGGAGAC 2700
Qy 2701 TCACTGTGCACTGAAGTTATGTATGACAGAGATTTTGGAGATGATGTCTCTATTC 2760
Db 2701 TCACTGTGCACTGAAGTTATGTATGACAGAGATTTTGGAGATGATGTCTCTATTC 2760
Qy 2761 AAGGGGGTGGGGGCGTTTTCAAATGTATGTCTGAGCACTGTGGATTGAGTCTCA 2820
Db 2761 AAGGGGGTGGGGGCGTTTTCAAATGTATGTCTGAGCACTGTGGATTGAGTCTCA 2820
Qy 2821 GTCCCTTCAACCAAGGCTGCGCACCTCTCATCTTGTGCTGTGCTGTGCTGTGCT 2880
Db 2821 GTCCCTTCAACCAAGGCTGCGCACCTCTCTCATCTTGTGCTGTGCTGTGCTGTGCT 2880
Qy 2881 AAAAAAAAAAAAAAAAAA 2902
Db 2881 AAAAAAAAAAAAAAAAAA 2902

RESULT 3
US-09-799-451-215
Sequence 215, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 215
LENGTH: 2501
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (106)..(1989)
US-09-799-451-215
Query Match 47.3%; Score 1372.2; DB 4; Length 2501;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTCCGACCCCACTTCCGCGTCCGCGACCACTGAGTCCGTTGCTTAC 141
DB 65 CTAATGATTCCTCCGCGCTTCTGCTCACTCTGCTGCGCACTGAGTCCGTTGCTTAC 124
QY 142 TCCAGCGCCGAGCGCAGAGTCTCCCTGCGCTTCCGCGCTGCGCTGCGAGAGCGCCGCGCC 201
DB 125 CGCGCGCTCCGCGCCCACTCCCTGCGCGCGAGCTA-----GCGCGCGCG 172
QY 202 TGGCGAGCGGAGCTCATGAGTCCCTTAACCTCTGATGAGAGACAGCGGCTGAGCGGC 261
DB 173 TGGCGAGAGGAGCTGATGAGTCCGCGAGCGCTTAATGAGAGACAGCGGCTGAGCGGC 232
QY 262 ACCATCAAAACACAACTGCGGACCGCTACAGATTCTGAGAGCGCTGCGGCAAGGCA 321
DB 233 ACCACCAAGACAACTGCGGACCGCTACAGATTCTGAGAGCGCTGCGGCAAGGCA 292
QY 322 CTTACGCGAAGTGAAGAGGACGAGAGAGCTCGGCGCGCTCTGCTGCGCTCATGATCCA 381
DB 293 CTTACGCGAAGTGAAGAGGACGCGGAGAGCTCGGCGCGCTCTGCTGCGCTCATGATCCA 352
QY 382 TCCGAAAGCAAAATGAAAGATGAGAGAGATCTGCTGCACTACCGAGAGAGATGAGA 441
DB 353 TCCGAAAGCAAAATGAAAGATGAGAGAGATCTGATGCACTACCGAGAGATGAGA 412
QY 442 TCAATGCTTCACTCAACCACTCCCACTCACTTGCATCTCAATGAGTCTTGAAGATGCA 501
DB 413 TCAATGCTTCACTCAACCACTCCCACTCACTTGCATCTCAATGAGTCTTGAAGATGCA 472
QY 502 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 473 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532
QY 562 AGCGGCGACCGCTGAGTGAAGCGGAGCGGAGCTTCTTCCGCAATGCTGCTGCTGCC 621
DB 533 AGCGGCGACCGCTGAGTGAAGCGGAGCGGAGCTTCTTCCGCAATGCTGCTGCTGCC 592
QY 622 TGCATCTGCTGCAACCACTCCCACTCACTTGCATCTCAATGAGTCTTGAAGATGCA 681
DB 593 TGCATCTGCTGCAACCACTCCCACTCACTTGCATCTCAATGAGTCTTGAAGATGCA 652
QY 682 TAGATGCAATGGAACATCAAGATGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCT 741
DB 653 TAGATGCAATGGAACATCAAGATGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCT 712
QY 742 GCAAGTCTCTGCAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
DB 713 GCAAGTCTCTGCAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
QY 802 GGAAGCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
DB 773 GGAAGCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832
QY 862 TGGTGCATGAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
DB 833 TGGTGCATGAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
QY 922 GTAAAGGAGCTTACCTGAGCGCGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
DB 893 GCAAGCGGCGCTTACCTGAGCGCGCACTTAACTCTGCTGCTGCTGCTGCTGCTGCTG 952
QY 982 TGTAAATGAGTGAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
DB 953 TGTAAATGAGTGAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
QY 1042 TCAAATGAGTGAACCCCACTGAGTCCGAGAGCAAGAGCCCTGCTGCTGCTGCTGCTG 1101
DB 1013 TCAAATGAGTGAACCCCACTGAGTCCGAGAGCAAGAGCCCTGCTGCTGCTGCTGCTG 1072
QY 1102 CTAGTGTGATCTTGGCGCGGCTTCACTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTG 1161
DB 1073 CTAGTGTGATCTTGGCGCGGCTTCACTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTG 1132

QY 1162 TCTGAGAGATGAGAGCCCAAGTGTGACGCTTCTTCAAGAGAGCACTGCTCCGAGAGTGA 1221
DB 1133 TCTGAGAGATGAGAGCCCAAGTGTGACGCTTCTTCAAGAGAGCACTGCTCCGAGAGTGA 1192
QY 1222 GCACTGATCTGAGGCTGAGAGCGGCACTTCTCTTGAAGAGTCCGAGAGAGAGATGAGA 1281
DB 1193 GCACTGATCTGAGGCTGAGAGCGGCACTTCTCTTGAAGAGTCCGAGAGAGAGATGAGA 1252
QY 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGAGATCTTCTGCTGCTGCTGCTGCTG 1341
DB 1253 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGAGATCTTCTGCTGCTGCTGCTGCTG 1312
QY 1342 GCTTAAAGCTTCCGAGAGGCACTTCTCAAGAGAGAGTCTTCTGCTGCTGCTGCTG 1401
DB 1313 ACTCAAGCTTCCGAGAGGCACTTCTCAAGAGAGAGTCTGCTGCTGCTGCTGCTGCTG 1372
QY 1402 AGAGAGAGCTTCAAGAGATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1461
DB 1373 AGAGAGAGCTTCAAGAGATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
QY 1462 TATCCCTGCTCCCAAGAGAGGCACTCTTAAAGATCTGCAAGCGTGAATCTGATCT 1521
DB 1427 ---CCTGCTCCCAAGAGAGGCACTTCTTCAAGAGAGGCACTGCTGCTGCTGCTGCT 1483
QY 1522 ACTCTCTCCAGAGCGGAGAGTCTGAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAG 1581
DB 1484 ACTCTCTCCAGAGCGGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1543
QY 1582 GTGGGAG 1641
DB 1544 GTGGGAG 1603
QY 1642 GCAATCTCAAACTCAATGAGAGAGTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701
DB 1604 GCAATCTCAAACTCAATGAGAGAGTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1663
QY 1702 CTTTGTGCTCCCTGAG 1761
DB 1664 CTTTGTGCTCCCTGAG 1723
QY 1762 CAGGAGCTGAG 1821
DB 1724 CAGGAGCTGAG 1783
QY 1822 CTGAAGCTCTTCCGAG 1881
DB 1784 CTGAAGCTCTTCCGAG 1843
QY 1882 TTGAGAGAGCTCCCTCAAGAG-----GTCTGAAGAGATGATGAGAGATCT 1929
DB 1844 TTGAGAGAGCTCCCTCAAG 1903
QY 1930 TGGGAGATGAGCTGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1989
DB 1904 TGGGAGATGAGCTGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1963
QY 1990 CCTTGAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2049
DB 1964 CACTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022
QY 2050 GCTCTGAGAGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2090
DB 2023 GCTCTGAG 2063

RESULT 4
US-09-949-016-2384
; Sequence 2384, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF


```

Db      824  CCAATGAGACTTAAAGGCAAGAAACCTGCTCTTGATGCTGATATGAAACATCAAGATTGC 883
Qy      710  TGAATTTGGCTCTTCCAACTGTACCAAAAGCAAGTTCTTCCAGAGCTTGTGAGAG 769
Db      884  AGACTTTGGCTTCCAGCATGATATCACCTTTGGGAAACAGCTGACACCTTCTGTGGCAG 943
Qy      770  CCCTCTAGGCGCTCGGCTGAGATGTCACAGGAAAGCCCTATGTGGCCCAAGAGTGA 829
Db      944  TCCCTCTTATGCTGCCCCCAAGACTCTTCCAGGCGCAAAAATATATGAGACCCGAGTGA 1003
Qy      830  CAGCTGTCTCTGAGCGCTTCTCTGTACATCTGTGATGACCAATGACCTTTTGAAG 889
Db      1004  TGTGTGAGGCTTGAAGATTAATCCTTAATACCTGTGACGCGAATCCCTGCTTTGATGG 1063
Qy      890  GCAAGATCATTAACACTGTGTGAAGCAATCATGACGGGGCTTACCGTGAAGCCCA 949
Db      1064  ACAGAACTCAAGAGGCTGCGGGAACGGGTACTGAGGGGAAATAACGGTATTCATTCTA 1123
Qy      950  GCGGTCCGATGCGCTGTG---GCTGATCCGCTGCTGTAAATGATGAACCCGACCGTGC 1006
Db      1124  CATGTCCACGACTGTGAAACCTGTCTTAAGAAATTTCTCATTTCTTAATCCACGAAAG 1183
Qy      1007  GGCCACACTGAGAGATGTAGCCAGTCAATGTGTGGGTCAACTGGGTTAC 1055
Db      1184  AGGACATTTAGAGCAATCATGAAGATCATGATGATGATGTGGTCAAC 1232

```

```

RESULT 8
US-09-984-890-1
; Sequence 1, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CI001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-890-1

```

```

Query Match      7.0%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 9.3e-42;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

Qy      290  CTACGAGTTCTCTGAGAGCGCTGGGCAAGGCACTTACGGGAGGTGAAGAGGACGAGA 349
Db      156  CTACCGGCTCTCAAGACCATTTGGCAAGGTAATTTTCCAGAGTGAAGTTGGCCGACA 215
Qy      350  GAGCTGGGGCGTCTGTGGCCATCAAGTCCATGAGAAAGCAAAATCAAAGTGAACA 409
Db      216  CATCTGACTGGGAAAGAGGTAGCTGTGAAGATCATTGACAAAGCTCACTGAATCTCTC 275
Qy      410  GGATCTGCTGCACATAGAGGAGGAGTTGAGATATGTCTTCACTCAACACCCCAAT 469
Db      276  CAGCTTCAGAAACTATTCGCGAAGTAAGAAATATAGAGTTTGAATCATCCCAAT 335
Qy      470  CATTGCCATCATGAAGTGTGAGATAGCAGCAAGATGTGATGTCTCATGAGATATGC 529
Db      336  AGTTAAATTAATTTGAAGTATGAGACTGAGAAACGCTCTAATCTGTCTATGAGTACGC 395
Qy      530  CAGCCGAGGAGATCTGTATATTAATCACTGAGAGCGGCAAGGCTGAGTGAAGCGGAGCG 589
Db      396  TAGTGGGAGAGAGTATTTGATTAATCACTGAGGCTCATGGGAGATGAAGAAAGAGGCG 455
Qy      590  CAGGCAATTTCTCGAAGATCGTGTCTGCCCTGCACTACTGCAACAGAAAGGAGATGCT 649

```

```

Db      456  TCGAGCCCAATTCGCGAGGATAGTGTCTGTGTCAGTATGTGACCAAGAGTTATTTG 515
Qy      650  TCAACGAGATCTCAAGCTGAAAAACATCTTCTAGATGCCAAATGAAACATCAAGATTGC 709
Db      516  CCAATGAGACTTAAGGCAAGAAACCTGCTCTTGATGCTGATATGAAATCAATCAAGATTGC 575
Qy      710  TGAATTTGGCTCTTCCAACTGTACCAAGGCAAGCAATTTCTCAAGAGCTTGTGGAG 769
Db      576  AGACTTTGGCTTCAAGCAATGAATTAACCTTTGGGAAACAGCTGACACCTTCTGTGGCAG 625
Qy      770  CCCTCTAGGCGCTCGCTGAGATGATCAAGGAAAGCCCTATGTGGGCCCAAGAGTGA 829
Db      636  TCCCTCTTATGCTGCCCCCAAGACTCTTCCAGGCGCAAAAATATATGAGAACCCAGAGTGA 695
Qy      830  CAGCTGTCTCTGAGCGCTTCTCTGTACATCTGTGATGACGACCAATGACCTTTTGAAG 889
Db      696  TGTGTGAGGCTTGAAGATTAATCCTTAATACCTGTGACGCGAATCCCTGCTTTGATGG 755
Qy      890  GCAAGATCATTAACACTGTGTGAAGCAATCATGACGGGGCTTACCGTGAAGCCGCA 949
Db      756  ACAGAACTCAAGAGGCTGCGGGAACGGGTACTGAGGGGAAATAACGGTATTCATTCTA 815
Qy      950  GCGGTCCGATGCGCTGTG---GCTGATCCGCTGCTGTAAATGATGAACCCGACCGTGC 1006
Db      816  CATGTCCACGACTGTGAAACCTGTCTTAAGAAATTTCTCATTTCTTAATCCACGAAAG 875
Qy      1007  GGCCACACTGAGAGATGTAGCCAGTCAATGTGTGGGTCAACTGGGTTAC 1055
Db      876  AGGACATTTAGAGCAATCATGAAGATCATGATGATGATGTGGTCAAC 924

```

```

RESULT 9
US-10-274-194-1
; Sequence 1, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CI001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-274-194-1

```

```

Query Match      7.0%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 9.3e-42;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

Qy      290  CTACGAGTTCTCTGAGAGCGCTGGGCAAGGCACTTACGGGAGGTGAAGAGGACGAGA 349
Db      156  CTACCGGCTCTCAAGACCATTTGGCAAGGTAATTTTCCAGAGTGAAGTTGGCCGACA 215
Qy      350  GAGCTGGGGCGTCTGTGGCCATCAAGTCCATGAGAAAGCAAAATCAAAGTGAACA 409
Db      216  CATCTGACTGGGAAAGAGGTAGCTGTGAAGATCATTGACAAAGCTCACTGAATCTCTC 275
Qy      410  GGATCTGCTGCACATAGAGGAGGAGTTGAGATATGTCTTCACTCAACACCCCAAT 469
Db      276  CAGCTTCAGAAACTATTCGCGAAGTAAGAAATATAGAGTTTGAATCATCCCAAT 335
Qy      470  CATTGCCATCATGAAGTGTGAGATAGCAGCAAGATGTGATGTCTCATGAGATATGC 529
Db      336  AGTTAAATTAATTTGAAGTATGAGACTGAGAAACGCTCTAATCTGTCTATGAGTACGC 395
Qy      530  CAGCCGAGGAGATCTGTATATTAATCACTGAGAGCGGCAAGGCTGAGTGAAGCGGAGCG 589

```



```
Db      396 TAGGCGGAGAGGATTTTAACTTAGTGGCTCATGCGAGATGAAAGAAAAGAGGC 455
Qy      590 CAGGATTTCTTCCGACAGATCGGTCTGCCCTGCACTAAGCCAGAGAGGATGCT 649
Db      456 TCGAGCCAAATTCGCGCAGAGTAGTGTCTGCTGCACTAAGTCAAGAAAGTTATGT 515
Qy      650 TCACCGAGATCTCAAGGTGAAAACATCTTCTAGATGCAATGAAAACATCAAGATTGC 709
Db      516 CCAATAGAGACTTAAAGGCAAGAAAACCTGCTCTTGATGCTGAATGAAACATCAAGATTGC 575
Qy      710 TGAATTTGGGCTCTCCCACTGTACCAAGGCAAGTTCTCCAGACGTTCTTGGGAG 769
Db      576 AGACTTTGGGCTCTCCCACTGTACCAAGGCAAGTTCTCCAGACGTTCTTGGGAG 635
Qy      770 CCGCTCTAGCGCTCGGCTGAGATAGTCAACGGAGAGCCCTATGTTGGGCGCCAGAGTGA 829
Db      636 TCCCCCTTATGCTGCCCCAGAACTCTTCAGGGCAAAAATATATGATGACCCGAGGTGA 695
Qy      830 CAGCTGCTCTGCGGCGCTTCTCTGTACATCTGTCATGTCATGTCAGCAATGCCCTTTGACG 889
Db      696 TGTGTGAGGCTAGGAGATTAATCTCTATACATGCTGTACAGCGGATCCCTGCTTTGATGG 755
Qy      890 GCAAGATCATPAAACACTGTGTAAAGCAATCATGTAACGGGCGTTACCGTGAGCGCCCA 949
Db      756 ACAGAACTCAAGAGAGCTGCGGGAACGGGTACTGAGGGGAAAATAACGTTATTCATTCTA 815
Qy      950 GCGGTCCGAGTCCGTG---GCCGATCCGCTGCTGTATGTTGTAAGAACCCCAACCGCTG 1006
Db      816 CATGTCAACGACTGTGAAAACCTGCTTAAGAAATTTCTCATCTTATTCACCAAGAG 875
Qy      1007 GGCCACACTGAGAGATGTAGCCACTGATGCTGCTGCTCAACTGCGGTTAC 1055
Db      876 AGGCACTTTAGAGCAATCATGAAGATTCATGATGATATGTTGGTAC 924
```

```
RESULT 10
US-09-270-767-1257/c
; Sequence 1257, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1257
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1257
```

Query Match 6.5%; Score 189.2; DB 4; Length 484;
Best Local Similarity 64.3%; Pred. No. 1.2e-38;
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

```
Qy      355 CGGGGCGTCTGTGGGCATCAAGTCCATCAGAAAACAAAATCAAAAGATGAGCAGATC 414
Db      442 CCGGCCAGAGAGTGGCTATCAAAACCATCAAGAGTGAAGATCGAGGCCAGCGGATTT 383
Qy      415 TGTGTCACATACGAGAGGAGATTTGATCTTCACTCAACACCCCACTCATTTG 474
Db      382 TGTGTCACATCCGTCGAGAGTGAATTAAGCTCAGTCAATCCCAACATCATCC 323
Qy      475 CCATCCATGAAGTGTGAGATAGCAGCAAGATTGATTTGATGAGATGATGCCAGCC 534
Db      322 ACATCTACGAAGTATTTGAAGATGTGAAGAAATGTGTCTAGTATGAAATTTCCCTG 263
Qy      535 GAGCGCATCTGTATGATTAATCATCACTGAGCGGCAACGCTGATGAGCGGAGCCAGCC 594
Db      262 GCGCGAGCTCTTACGATATCTGTCTGAAAGAAAGTTCTCACACGAGAGAGCGAGAC 203
```

```
Qy      595 ATTTCTCCGACAGATCGTCTTGGCCTGCACTAATGCAACAGAACGGATGCTTACC 654
Db      202 GCATCTTCCGCGCAGAGTGGCCACCGCGCTTACTAATGTCAACAGCAAGATCTGCCATC 143
Qy      655 GAAATCTCAGCTGGAAGAAACATCTTCTAATGCAATGGAACATCAAGATTGCTGACT 714
Db      142 GCAATCTCAAGCTGGAAGAAACATCTTCTGAGCAGAAAGGCAATGCTAAGATTGCTAAT 83
Qy      715 TTGGCCTTCCCAACTGTACCAAGAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTC 774
Db      82 TTGGGTTGTGAATGTGTTGATGACCAAGGACTGTGGGACCTTTTGGGGTTCCAC 23
Qy      775 TCTACGCTTCGCTGAGATAGT 796
Db      22 TCTATGCTTCGCGGAAATTTGT 1
```

```
RESULT 11
US-09-270-767-16539/c
; Sequence 16539, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16539
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16539
```

Query Match 6.5%; Score 189.2; DB 4; Length 484;
Best Local Similarity 64.3%; Pred. No. 1.2e-38;
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

```
Qy      355 CGGGGCGTCTGTGGGCATCAAGTCCATCAGAAAACAAAATCAAAAGATGAGCAGATC 414
Db      442 CCGGCCAGAGAGTGGCTATCAAAACCATCAAGAGTGAAGATCGAGGCCAGCGGATTT 383
Qy      415 TGTGTCACATACGAGAGGAGATTTGATCTTCACTCAACACCCCACTCATTTG 474
Db      382 TGTGTCACATCCGTCGAGAGTGAATTAAGCTCAGTCAATCCCAACATCATCC 323
Qy      475 CCATCCATGAAGTGTGAGATAGCAGCAAGATTGATTTGATGATGATGATGATGATG 534
Db      322 ACATCTACGAAGTATTTGAAGATGTGAAGAAATGTGTCTAGTATGAAATTTGCCCTG 263
Qy      535 GAGCGCATCTGTATGATTAATCATCACTGAGCGGCAACGCTGATGAGCGGAGCCAGCC 594
Db      262 GCGCGAGCTCTTACGATATCTGTCTGAAAGAAAGTTCTCACACGAGAGAGCGAGAC 203
Qy      595 ATTTCTCCGACAGATCGTCTTGGCCTGCACTAATGCAACAGAAAGGATGCTTAC 654
Db      202 GCATCTTCCGCGCAGAGTGGCCACCGCGCTCTACTACTGTCAACAGCAAGATGTGCCATC 143
Qy      655 GAAATCTCAGCTGGAAGAAACATCTTCTAATGCAATGGAACATCAAGATTGCTGACT 714
Db      142 GCAATCTCAAGCTGGAAGAAACATCTTCTGAGCAGAAAGGCAATGCTAAGATTGCTAAT 83
Qy      715 TTGGCCTTCCCAACTGTACCAAGAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTC 774
Db      82 TTGGGTTGTGAATGTGTTGATGACCAAGGACTGTGGGACCTTTTGGGGTTCCAC 23
Qy      775 TCTACGCTTCGCTGAGATAGT 796
Db      22 TCTATGCTTCGCGGAAATTTGT 1
```

RESULT 12
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 v1
; CURRENT APPLICATION NUMBER: US/09/930.181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match 6.4%; Score 186.4; DB 3; Length 2908;
Best Local Similarity 53.5%; Pred. No. 1.7e-37;
Matches 413; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

```
QY 288 CGCTACGAGTTCTGAGAGCGCTGGCAAGGCGACCTACGGGAAGTGAGAGGACGACGA 347
DB 157 CCTACCGGCTGAGAGAGCGCTGGCAAGGCGACGAGTCTGTGTAAGCTGGGGGTT 216
QY 348 GAGAGCTCGGGGCGCTGTGTGGCCATCAATCCAGAAAGCAAAATCAAGATGAG 407
DB 217 CACGCGCTCACCTCCAGAGGTGGCCATCAAGTTCGCAACCGTAGAGACTGAGGAG 276
QY 408 CAGAGTCTGTGCACATACGAGGAGATTGAGATCAATGTCTTCACTCAACACCCGAC 467
DB 277 TCGGTGCTGATGAGGTGAGCGGAGATCGCATCTCGAAGCTCATTAAGCACCCGAC 336
QY 468 ATCATTCGCATCCATGAAAGTTTGAAGATAGCGAAGATTGTGATTTGATGAGAT 527
DB 337 GTCTTAAGCTGACGAGCGTTTATGAAACAAATAATTTGTACTGTGCTTGAACAC 396
QY 528 GCCAGCGGAGCGATCTGTATGATTAATCATGAGCGGCGACGAGTGAAGCGGAGAC 587
DB 397 GTGTACAGGTGTGAGCTCTTCTGACTACTGTGTAAGAGGAGGCTGACCGCTTAAG 456
QY 588 GCCAGGATTTCTTCCGACAGATGTGTGCTGCTGCACTACTGCAACAGAGCGGATC 647
DB 457 GCTCGAAGTTCTTCCGGCAGATCATCTGTGCTGCACTTCTGCAACGCACTTCATA 516
QY 648 GTTACCGAGATCTCAAGCTGGAAAAACATCTTTAGATGCCAATGGAACATCAAGAT 707
DB 517 TGCCACAGGATCTGAAACCTGMAAACCTCTGTGACGAGAAACAACATCCGATC 576
QY 708 GCTGACTTTGGCCCTCTCAACCTGTACCAAAAGGCAAGTCTCCGAGCGTTGTGGG 767
DB 577 GCAGACTTTGGCATGGGCTCCCTCAAGTTGGCAAGCCCTGTGGAGACAGCTGGGG 636
QY 768 AGCCCTCTCAAGCCTGCGCTGAGATAGTCAACGGGAAGCCCTTATGTGGGCCAGAG 827
DB 637 TCCCCCACTACGCTGCGCCGAGGTGATCCGGGGGAGAAATGACGGCCGGAAGGCG 696
QY 828 GACAGCTGTCTCTGGGCGTTCTCTGTATCATCTGTGTGATGGCACCACATGCCCTTT 887
DB 697 GACCTGTGTGAGCTGCGGCGTCACTCTGTCCTGTGTGCTGTGGGGGCTGTGCCCTT 756
QY 888 GGGGAGGATCATATAAAGCTGTGTGAGCAAGATCAAGTAACGGGGGCTTACCGTGAAG 947
DB 757 GATGACAACTTGGCAGAGCTGTGTGAGAGGTGAAGCGGGGCGGTTCACATCCGAC 816
QY 948 ---AAGCGTTCGATGCTGTGTGCTGATCCGATGCTGTATATGTGAACCCACCGCT 1004
DB 817 TTATATCCGCGGCACTGCAAGATCTGTGTGCGGAGCATGATGAGGTGAGCGCCGACGC 876
```

RESULT 13
US-08-557-006C-39
; Sequence 39, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Bert, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557.006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 2652
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
US-08-557-006C-39

Query Match 6.3%; Score 182.4; DB 3; Length 2652;
Best Local Similarity 53.0%; Pred. No. 1.7e-36;
Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;

```
QY 242 GAGAGCGGCGTGAAGCGGACGATCAACAAACACCTGCGGACCGCTACGAGTTCT 301
DB 4 GAACATGCTGAGAAAGCAAGACGAGCGGCGTGTGAATCGACATCACTGTGCGG 63
QY 302 GAGAGCGCTGGGACGAGGCACTACGGAAGTGAAGAGCGAGAGCTCG---GG 358
DB 64 GACACCTCTGGGCGTGGCACTTCGCGAAGTGAAGATTGGAACATCAATTGACAGG 123
QY 359 GCGTCTGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGCAGATCTGCT 418
DB 124 CCAATAAGTGGCAGTTAAGATCTTAATATAGACAGAAATTCGAGTTAAGTGTGG 183
QY 419 GCACATACGAGGAGATTAAGATCATGTCTTCACTCAACACCCCACTCATTTGCAAT 478
DB 184 AAAAATAAAGAGAAATTCATAAATCTTTGCTGATCCCTCATTTATCAAACT 243
QY 479 CCAATGAAGTTTGAAGATAGCAAGATTTGATTTGATGATGATGATGATGATGATGAT 538
DB 244 CTACCAAGATATACGACTCCAAACAGACTTTTATATGATATGATATGATGATGATG 303
QY 539 CGATCTGTATGATTACATCAAGTGAAGCGGCAACGCGTGAAGTGAAGCGGACGCGCAT 598
DB 304 TGAATTTGTCATCATCTGTAAACGCGGAGGTTGAAGAGTGAAGTGAAGTGAAGTGA 363
QY 599 CTTCGACAGATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
DB 364 CTTCGACAGATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
QY 659 TCTCAAGCTGGAAGAACATCTCTTCAAGTGCATATGGAACATCAAGATTTGCTGCT 718
DB 424 CTGAAAGCAGAGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2005, 22:57:45 ; Search time 1510.61 Seconds
(without alignments)
11372.259 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902
Sequence: 1 cactagtgtgataccaagaat.....aaaaaaaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: geneseq1980s:*
3: geneseq1990s:*
4: geneseq2000s:*
5: geneseq2001as:*
6: geneseq2001bs:*
7: geneseq2002as:*
8: geneseq2002bs:*
9: geneseq2003as:*
10: geneseq2003bs:*
11: geneseq2003cs:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2902	100.0	2902	4	AAC90433
2	2863.4	98.7	3073	13	ABD33083
3	2066.8	71.2	2929	6	ABD31710
4	1938.8	66.8	37278	13	ABD33082
5	1379	47.5	3353	6	ABK14000
6	1375.8	47.4	3395	12	ADL14160
7	1375.8	47.4	3395	12	AAK94280
8	1375.8	47.4	3395	12	ADL30885
9	1375.8	47.4	3404	13	ABD33085
10	1374.2	47.4	3443	5	ABX71420
11	1374.2	47.4	3443	10	ADF76964
12	1374.2	47.4	3443	10	ADF81952
13	1374.2	47.4	3443	12	ADP020171
14	1374.2	47.4	3463	12	ADJ96554
15	1372.2	47.3	2501	12	ADMA3851
16	1371.8	47.2	2043	6	ABZ11333
17	1370.6	47.2	3360	6	ABD26459
18	1360.2	46.9	1884	12	ADL1162
19	1349.2	46.5	2291	4	AAK44659
20	1349.2	46.5	2291	12	AD129357

21 1323.2 45.6 3200 4 AAF75338
22 1230 42.4 1833 10 ABZ77163
23 849.2 29.3 2616 6 ABQ72599
24 849.2 29.3 2619 6 ABQ72698
25 732 25.2 39699 -13 ABD33084
26 507.6 17.5 6828 4 AAD14327
27 507.6 17.5 6828 10 ADE38420
28 507.6 17.5 6828 11 ADN95767
29 507.6 17.5 6828 12 ADL74808
30 507.6 17.5 6828 12 ADL25352
31 507.6 17.5 6828 12 ADL9733
32 507.6 17.5 6828 12 ADP43253
33 507.6 17.5 6828 13 ADR25680
34 507.6 17.5 6854 12 AADQ23883
35 506.4 17.5 2884 4 AAD14328
36 360.2 12.4 587 4 AAK93296
37 360.2 12.4 587 4 AAK91887
38 360.2 12.4 587 12 ADL29723
39 360.2 12.4 3594 5 AAD03994
40 356.8 12.3 1454 4 ABL21269
41 296.8 10.2 1454 4 ABL1802
42 268.2 9.2 1723 4 AAF44655
43 236.6 8.2 1594 8 AAF60326
44 236.6 8.2 1594 12 AD129353
45 236.6 8.2 1594 12 AD129353

ALIGNMENTS

RESULT 1
AAC90433
ID AAC90433 standard; cDNA; 2902 BP.
XX
AC AAC90433;
XX
DT 19-MAR-2001 (first entry)
XX
DE Murine lymph node Stromal cell kinase 1 coding sequence.
XX
KW Murine; Lymph node Stromal cell kinase; MSK-1; autoimmune disorder;
KW wound healing; periodontal disease; inflammatory disease; tumour;
KW infection; allergy; ss.
XX
XX OS Mus musculus.
XX
XX WO200073468-A1.
XX
PD 07-DEC-2000.
XX
PP 26-MAY-2000; 2000WO-US014696.
XX
PR 28-MAY-1999; 99US-0136781P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Bird TR, Virca GD, Martin U, Anderson DM;
XX
DR WPI; 2001-061546/07.
XX
DR P-PSDB; AAB50056.
XX
PT Novel murine and human kinase nucleic acids useful for treating
PT inflammations, infections, tumors, allergies, autoimmune diseases, and
PT for stimulating or suppressing immune responses.
XX
PS Claim 1; Page 86-87; 106pp; English.
XX
The present sequence is the coding sequence for Murine lymph node Stromal
cell kinase 1 (MSK-1). The protein encoded by the present sequence is
useful for treating a variety of disorders listed in the disclosure of
the specification, including autoimmune disorders, allergic reactions,
myeloid or lymphoid cell deficiencies, wound healing and tissue repair
and replacement, burns, incisions and ulcers, periodontal disease,

CC Inflammatory diseases, tumours and bacterial, viral or fungal infection
XX Sequence 2902 BP; 654 A; 794 C; 796 G; 658 T; 0 U; 0 Other;
Query Match 100.0%; Score 2902; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACTAGTGGATCCAAAGATTGGCAGAGGCTGCTGGGTGGGCTGTGACTGTGAG 60
DB 1 CACTAGTGGATCCAAAGATTGGCAGAGGCTGCTGGGTGGGCTGTGACTGTGAG 60
QY 61 CCCGGGCTCAGCGCGGCTGTACTGTGCTGCCGACCCATCCACTCCCTCGGGTCCCCGA 120
DB 61 CCCGGGCTCAGCGCGGCTGTACTGTGCTGCCGACCCATCCACTCCCTCGGGTCCCCGA 120
QY 121 CCAATGAGTGGGTGGCTTACTCCAGGCCCCGAGCCAGGCTCCCTCGGCTCCGCTGG 180
DB 121 CCAATGAGTGGGTGGCTTACTCCAGGCCCCGAGCCAGGCTCCCTCGGCTCCGCTGG 180
QY 181 CTTGGAGAGGCGCCGCGCGCTGGCGGACGGGCTCATCAAGTCCCTTAACCTTGATGA 240
DB 181 CTTGGAGAGGCGCCGCGCGCTGGCGGACGGGCTCATCAAGTCCCTTAACCTTGATGA 240
QY 241 AGAAGCAGGGGGTGAAGGCGCATCATCAAAACAACCTGGGCGACCGTACAGAGTTCC 300
DB 241 AGAAGCAGGGGGTGAAGGCGCATCATCAAAACAACCTGGGCGACCGTACAGAGTTCC 300
QY 301 TGGAGCGCTGGGCGAGGCGCATTCACGGGAAAGTGAAGAAAGCAGAGAGCTCGGGGC 360
DB 301 TGGAGCGCTGGGCGAGGCGCATTCACGGGAAAGTGAAGAAAGCAGAGAGCTCGGGGC 360
QY 361 GTCTGGTGGCCATCAAGTCCATCAGAAAGAACAATTCAAAGATGAGCAGGATGTCTGC 420
DB 361 GTCTGGTGGCCATCAAGTCCATCAGAAAGAACAATTCAAAGATGAGCAGGATGTCTGC 420
QY 421 ACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTCGCCATCC 480
DB 421 ACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTCGCCATCC 480
QY 481 ATGAAGTGTGAGAAATAGCAGCAAGATTGATGTTCATGAGATATGCCAGCGAGGCG 540
DB 481 ATGAAGTGTGAGAAATAGCAGCAAGATTGATGTTCATGAGATATGCCAGCGAGGCG 540
QY 541 ATCTGATGATTAATCATCAATGAGCGGCCACGGCTGAGTGAAGCGGAGCGCCAGCATTTCT 600
DB 541 ATCTGATGATTAATCATCAATGAGCGGCCACGGCTGAGTGAAGCGGAGCGCCAGCATTTCT 600
QY 601 TCCGACGATCGGTCTGCCCTGCACTACGCGCACAGAACGGGATGTTCAACGGAATC 660
DB 601 TCCGACGATCGGTCTGCCCTGCACTACGCGCACAGAACGGGATGTTCAACGGAATC 660
QY 661 TCAAGCTGAAAAATCATCTTCTAGATGCAATGAGAAATCAAGATTGCTGACTTTGGCC 720
DB 661 TCAAGCTGAAAAATCATCTTCTAGATGCAATGAGAAATCAAGATTGCTGACTTTGGCC 720
QY 721 TCTTCAAACCTGTACCAAAAGCAAGTTCTCCAGAGTTCGTGGAGGCCCTCTACAG 780
DB 721 TCTTCAAACCTGTACCAAAAGCAAGTTCTCCAGAGTTCGTGGAGGCCCTCTCTACAG 780
QY 781 CTTGGAGAGTGAAGTGAAGGAGGAGGCTTATGAGGCGCCAGAGTGAAGAGCTGTGCTC 840
DB 781 CTTGGAGAGTGAAGTGAAGGAGGAGGCTTATGAGGCGCCAGAGTGAAGAGCTGTGCTC 840
QY 841 TGGGCGTTCTCTGTATCATCTGTGATGAGCAACATGCTTGAAGGGAGGATCATTA 900
DB 841 TGGGCGTTCTCTGTATCATCTGTGATGAGCAACATGCTTGAAGGGAGGATCATTA 900
QY 901 AAAACCTGTGAAGCAAAATCAAGTAACGGGGCTTACCGTGAAGCCGCCAAGCGCTCGATG 960
DB 901 AAAACCTGTGAAGCAAAATCAAGTAACGGGGCTTACCGTGAAGCCGCCAAGCGCTCGATG 960
QY 961 CTTGTGGCTGATCCGGTGGCTTTAATGATGAACCCCAACCGTGGGCCACACTGAGAG 1020

DB 961 CTTGTGGCTGATCCGGTGGCTTTAATGATGAACCCCAACCGTGGGCCACACTGAGAG 1020
QY 1021 ATGAGCCAGTCAATGTGTGGGTCAACTGGGATTAACACACCGAGTGGGGAACAGAGAG 1080
DB 1021 ATGAGCCAGTCAATGTGTGGGTCAACTGGGATTAACACACCGAGTGGGGAACAGAGAG 1080
QY 1081 CCTGTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCTCCATGGCGGACTGTGT 1140
DB 1081 CCTGTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCTCCATGGCGGACTGTGT 1140
QY 1141 TAGGTGCTCTCGCGCGCCCTCCTGAGAGATGAGCAAGTGTGAGCTTCTTAAGC 1200
DB 1141 TAGGTGCTCTCGCGCGCCCTCCTGAGAGATGAGCAAGTGTGAGCTTCTTAAGC 1200
QY 1201 AGCAGTGGCGGAGGTGAAGACCTGTACCTGGGCTGAGCGGCAACATTTCTTAAGA 1260
DB 1201 AGCAGTGGCGGAGGTGAAGACCTGTACCTGGGCTGAGCGGCAACATTTCTTAAGA 1260
QY 1261 AGTCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTACCCGGCTGAGGATACCT 1320
DB 1261 AGTCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTACCCGGCTGAGGATACCT 1320
QY 1321 CTTCTGGCCCTGGCAAGAGAGGCTTAAGCTTCGGAAGGCAATTCGAAGAAAAGTCT 1380
DB 1321 CTTCTGGCCCTGGCAAGAGAGGCTTAAGCTTCGGAAGGCAATTCGAAGAAAAGTCT 1380
QY 1381 CTACCTGTGAGGAGAGTACAGAGAGACCTCAGAACTCAGACCGGCTGATATCTC 1440
DB 1381 CTACCTGTGAGGAGAGTACAGAGAGACCTCAGAACTCAGACCGGCTGATATCTC 1440
QY 1441 CAGGCGAGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAGGATCTTAAGAGTCTC 1500
DB 1441 CAGGCGAGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAGGATCTTAAGAGTCTC 1500
QY 1501 GACAGGATGATGTGGTTACTACTCTCTCCAGAGCCCAAGGCTGTGGGAACTCTAG 1560
DB 1501 GACAGGATGATGTGGTTACTACTCTCTCCAGAGCCCAAGGCTGTGGGAACTCTAG 1560
QY 1561 ACGCAGATGATGTGTTGTGAGTGGGAGCCCGTGAAGCAAGTCTCCACAGGCTTCAG 1620
DB 1561 ACGCAGATGATGTGTTGTGAGTGGGAGCCCGTGAAGCAAGTCTCCACAGGCTTCAG 1620
QY 1621 GGTCTCTCTCCACCGCAAGGCAATTCCTCAATGCAATGCAAGTTCCTCCGCAACGCT 1680
DB 1621 GGTCTCTCTCCACCGCAAGGCAATTCCTCAATGCAATGCAAGTTCCTCCGCAACGCT 1680
QY 1681 TAGAAGCATTACCCCTAGACCTTTGGGCTCCCTGAGCAAACTGGGCTCTCCCATCTG 1740
DB 1681 TAGAAGCATTACCCCTAGACCTTTGGGCTCCCTGAGCAAACTGGGCTCTCCCATCTG 1740
QY 1741 CAGCCGCGCCAGCCGCTCAGGGGCTGTGAGTGAAGACAGATCTGCTCCGAGT 1800
DB 1741 CAGCCGCGCCAGCCGCTCAGGGGCTGTGAGTGAAGACAGATCTGCTCCGAGT 1800
QY 1801 CTTTGAACAATTGACTTGTCTGAAAGTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
DB 1801 CTTTGAACAATTGACTTGTCTGAAAGTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
QY 1861 CTGTGACCAACTGAGGGGGCTTGAAGAGCTCTCCCTCAAGAGGTGTGAACGATGTGGC 1920
DB 1861 CTGTGACCAACTGAGGGGGCTTGAAGAGCTCTCCCTCAAGAGGTGTGAACGATGTGGC 1920
QY 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTGACAGACTGCCAAGAGGTGACTGACGCT 1980
DB 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTGACAGACTGCCAAGAGGTGACTGACGCT 1980
QY 1981 ACAAGCAAGCCTTAGAATCTGCTCAAAGCTCAGCTGAGAGAGAGAGATGTGCTTAGT 2040
DB 1981 ACAAGCAAGCCTTAGAATCTGCTCAAAGCTCAGCTGAGAGAGAGAGATGTGCTTAGT 2040
QY 2041 ATGGGGTAGGCTGTGAGAGGTTTTCAGAGAAACCTGTGGGTGGATTCCTCACTGAATA 2100

Db 2041 ATGGGAGGAGCTCTGAGAGGGTTTGCAGAGAACTTGGGAGTTCCTCAGTAATA 2100
 Qy 2101 GAGTACATCAAGGGCTCTACGTCTGACGCTGACCTGAACCTGAAGAATGAGAAATCGC 2160
 Db 2101 GAGTACATCAAGGGCTCTACGTCTGACGCTGACCTGAACCTGAAGAATGAGAAATCGC 2160
 Qy 2161 ATTGATGTGAAAGAAATGGAAACCTTGTCTGCCGAGTGTATATAGTGGGCTCTGAA 2220
 Db 2161 ATTGATGTGAAAGAAATGGAAACCTTGTCTGCCGAGTGTATATAGTGGGCTCTGAA 2220
 Qy 2221 GGGGCTTACCTCTCTTGTGCAATGAGTCAACCCATGACATTTCCACCTCTCTCTCG 2280
 Db 2221 GGGGCTTACCTCTCTTGTGCAATGAGTCAACCCATGACATTTCCACCTCTCTCTCG 2280
 Qy 2281 CTGCACTTCAATTAAGTTTCTGTTTCAATCAACCAAGGGTATGAAACCTGACTTCT 2340
 Db 2281 CTGCACTTCAATTAAGTTTCTGTTTCAATCAACCAAGGGTATGAAACCTGACTTCT 2340
 Qy 2341 GGGAGGTAATGTGATGACTGCTCCATTAATTAAGAGAAACAGCTCTGCTTCCATCT 2400
 Db 2341 GGGAGGTAATGTGATGACTGCTCCATTAATTAAGAGAAACAGCTCTGCTTCCATCT 2400
 Qy 2401 CTGCTGTGTGATCTTCAAGACCTGGGAAGACTCGGACCGCTGTGACTTCAATCCTCA 2460
 Db 2401 CTGCTGTGTGATCTTCAAGACCTGGGAAGACTCGGACCGCTGTGACTTCAATCCTCA 2460
 Qy 2461 GGGGACCAAGATGCCCCCTGAGACCCCATCTTGAATCTGAGAGCTTGAAGCTGTT 2520
 Db 2461 GGGGACCAAGATGCCCCCTGAGACCCCATCTTGAATCTGAGAGCTTGAAGCTGTT 2520
 Qy 2521 CCTAGTACCAAGATGTGATGATGATCTCTGTTCTCAGGCCAACGGGACCTAGAAATGTC 2580
 Db 2521 CCTAGTACCAAGATGTGATGATGATCTCTGTTCTCAGGCCAACGGGACCTAGAAATGTC 2580
 Qy 2581 TGACTTATTAATTTTGTGATCTCACTTCTGTTTGTGTTTGTGTTTGTGTTG 2640
 Db 2581 TGACTTATTAATTTTGTGATCTCACTTCTGTTTGTGTTTGTGTTTGTGTTG 2640
 Qy 2641 TTTTGTGTTTAAATGTAATTTGCTGCTTCAATTAATGTAATGCTGTTTGGGGAAC 2700
 Db 2641 TTTTGTGTTTAAATGTAATTTGCTGCTTCAATTAATGTAATGCTGTTTGGGGAAC 2700
 Qy 2701 TCCACTGTGCACTGAAGTTTATGATACAGAAATATTTGGCAATGATGTCCTCTATTC 2760
 Db 2701 TCCACTGTGCACTGAAGTTTATGATACAGAAATATTTGGCAATGATGTCCTCTATTC 2760
 Qy 2761 AAGGGGGGTGGGGGCTTTTCAATGATGCTTGAAGCACTGTCTGGAATGAGTCTCA 2820
 Db 2761 AAGGGGGGTGGGGGCTTTTCAATGATGCTTGAAGCACTGTCTGGAATGAGTCTCA 2820
 Qy 2821 GTCCCTTCAACCAAGGCTGCGCAACCTCCATCTTCAATCTGTGCGCAAAAAA 2880
 Db 2821 GTCCCTTCAACCAAGGCTGCGCAACCTCCATCTTCAATCTGTGCGCAAAAAA 2880
 Qy 2881 AAAAAAAAAAAAAAAAAA 2902
 Db 2881 AAAAAAAAAAAAAAAAAA 2902
 RESULT 2
 ID ABD33083 standard; cDNA; 3073 BP.
 AC ABD33083;
 XX 18-NOV-2004 (first entry)
 XX Murine cancer-associated (CA) cDNA MR07-004.
 XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW 86; cancer; cytostatic.
 OS Mus musculus.

XX MO2004058146-A2.
 XX 15-JUL-2004.
 XX 15-DEC-2003; 2003MO-US040081.
 XX 17-DEC-2002; 2002US-00322281.
 XX (SAGR-) SAGRES DISCOVERY INC.
 XX Morris DW, Malandro MS;
 XX WPI; 2004-499109/47.
 XX P-PSDB; AB084756.
 PT Novel human cancer associated protein encoded within open reading frame
 PT of cancer associated gene, useful as targets for diagnosing cancer.
 PS Disclosure; SEQ ID NO 22; 182bp; English.
 CC The invention relates to cancer-associated proteins (CAP) and the cancer-
 CC associated (CA) nucleic acids encoding them. The invention also relates
 CC to a method for treating cancers involving administering to a patient an
 CC inhibitor of CAP, and a method of screening for anticancer activity in a
 CC potential drug involving providing a cell that expresses a CA gene,
 CC contacting a tissue sample derived from a cancer cell with an anticancer
 CC drug candidate and monitoring the effect of the anticancer drug candidate
 CC on expression of the CA gene. The CAP proteins are useful for detecting
 CC cancer associated with expression of a CAP protein in a test cell sample
 CC and for screening for a bioactive agent capable of modulating the
 CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 CC cancer. Involving determining the expression of a CA nucleic acid in a
 CC tissue. This sequence represents murine CA cDNA of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3073 BP; 690 A; 827 C; 844 G; 712 T; 0 U; 0 Other;
 Query Match 98.7%; Score 2863.4; DB 13; Length 3073;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2864; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 32 CGTGTGGGTGGGGCTGTGACCTGTGAGCCCGGCTCAGCGCGCTGACTATGCTGC 91
 Db 1 GGTGTGGGTGGGGCTGTGACCTGTGAGCCCGGCTCAGCGCGCTGACTATGCTGC 60
 Qy 92 CCGACCACTTCACTTCCGCGGTCCCGCACATGATGAGTGGCTTACTCCAGCGCC 151
 Db 61 CCGACCACTTCACTTCCGCGGTCCCGCACATGATGAGTGGCTTACTCCAGCGCC 120
 Qy 152 GAGCCAGGCTTCCCTGCGCTCCGCTGCGCTTGGAGAGCGCCGCGCTGCGGACG 211
 Db 121 GAGCCAGGCTTCCCTGCGCTCCGCTGCGCTTGGAGAGCGCCGCGCTGCGGACG 180
 Qy 212 GCTCATTAAGTCCCTTAAACCTCTGATGAAGAGCGCGGTGAAGGGACCACTTCACA 271
 Db 181 GCTCATTAAGTCCCTTAAACCTCTGATGAAGAGCGCGGTGAAGGGACCACTTCACA 240
 Qy 272 ACACAACTTCCGCGCTACGAGTTCTGTGAGACGCTGGGCAAGGGCACTTACGAGAA 331
 Db 241 ACACAACTTCCGCGCTACGAGTTCTGTGAGACGCTGGGCAAGGGCACTTACGAGAA 300
 Qy 332 GGTGAAGAGGCAAGAGAGCTTGGGGCTGTGGTGGCCATCAAGTCCATGAGAAAGA 391
 Db 301 GGTGAAGAGGCAAGAGAGCTTGGGGCTGTGGTGGCCATCAAGTCCATGAGAAAGA 360
 Qy 392 CAATAATCAAGAGAGAGATCTGTGCAATTCAGGAGGAGGATTAAGATCATGCTTC 451
 Db 361 CAATAATCAAGAGAGAGATCTGTGCAATTCAGGAGGAGGATTAAGATCATGCTTC 420
 Qy 452 ACTCAACCAACCCCACTATTCATTCATGAAGTGTGAGATAGAGCAAGATTGT 511

Db 421 ACTCAACCAACCCCACTCATCTTGGCCATTCATGAAAGTGTGAGAAATAGCAAGCAATTTGT 480
QY 512 GATTGTGATGAGATATSCAGCCGAGGCGATCTGTATGATTAATCATCATGAGCGCCACG 571
Db 481 GATTGTGATGAGATATSCAGCCGAGGCGATCTGTATGATTAATCATCATGAGCGCCACG 540
QY 572 GCTGAGTGAACGGGAGCCGAGGCAATTTCTTCCGACAAATGTGTCTGCTCCGACATCTG 631
Db 541 GCTGAGTGAACGGGAGCCGAGGCAATTTCTTCCGACAAATGTGTCTGCTCCGACATCTG 600
QY 632 CCAACGAGAAGGGATCGTTACCCGAGATCTCAAGCTGGAAGAAATCTCTTCTAGATGCGAA 691
Db 601 CCAACGAGAAGGGATCGTTACCCGAGATCTCAAGCTGGAAGAAATCTCTTCTAGATGCGAA 660
QY 692 TGGAAAATCAAGATTTGCTGAATTTGGCCCTCTCCAACTGTACCAAGCAAGCAATTTCT 751
Db 661 TGGAAAATCAAGATTTGCTGAATTTGGCCCTCTCCAACTGTACCAAGCAAGCAATTTCT 720
QY 752 CCAAGACGTTCTGTGGAGAGCCCTCTTACGCTCGCTGAGATATGTAACGAGAAAGCCCTTA 811
Db 721 CCAAGACGTTCTGTGGAGAGCCCTCTTACGCTCGCTGAGATATGTAACGAGAAAGCCCTTA 780
QY 812 TGTGGAGCCCAAGAGGTGAGACAGTGTCTCTGGAGGCTTCTCCTGTATCAATCCTGTGTGATGG 871
Db 781 TGTGGAGCCCAAGAGGTGAGACAGTGTCTCTGGAGGCTTCTCCTGTATCAATCCTGTGTGATGG 840
QY 872 CACCATGCGCTTTGACGAGGAGATCATATAAACCTGTGAGAAATCAAGTAAACGAGGCG 931
Db 841 CACCATGCGCTTTGACGAGGAGATCATATAAACCTGTGAGAAATCAAGTAAACGAGGCG 900
QY 932 TTACCCGTGAACCCGCGCAAGCCGTCCGATGCTGTGGCTGTATCCGGTGGCTGTAAATGCT 991
Db 901 TTACCCGTGAACCCGCGCAAGCCGTCCGATGCTGTGGCTGTATCCGGTGGCTGTAAATGCT 960
QY 992 GAACCCCAACCCGCTGAGGCGACACTGAGAGATGTAGCCAGTCAATGATGGGCTCAACTGAGG 1051
Db 961 GAACCCCAACCCGCTGAGGCGACACTGAGAGATGTAGCCAGTCAATGATGGGCTCAACTGAGG 1020
QY 1052 TTACACCAACCGAGATCGAGGAAACAGAAAGCCCTGCGTGAAGGTGGGACCCCTAGTGTGTA 1111
Db 1021 TTACACCAACCGAGATCGAGGAAACAGAAAGCCCTGCGTGAAGGTGGGACCCCTAGTGTGTA 1080
QY 1112 CTTTGGCGCGGCGCTTCAATGCGGAGACTGTATACGTGCTCTCTCGCGCCCTCTCTGAGAA 1171
Db 1081 CTTTGGCGCGGCGCTTCAATGCGGAGACTGTATACGTGCTCTCTCGCGCCCTCTCTGAGAA 1140
QY 1172 TGGAGCCAAAGGTGTGACAGCTTCTTCAAGACAGACGTGCGGAGAGGTGGAAGCACTGTACC 1231
Db 1141 TGGAGCCAAAGGTGTGACAGCTTCTTCAAGACAGACGTGCGGAGAGGTGGAAGCACTGTACC 1200
QY 1232 TGGGCTGAGACCGGCAATTTCTTTAAAGATCCCGAAAGAGATGACATGTGCTCAAA 1291
Db 1201 TGGGCTGAGACCGGCAATTTCTTTAAAGATCCCGAAAGAGATGACATGTGCTCAAA 1260
QY 1292 TCTGCAAGGTGACCCCGGCTGAGATACCTTCTTCCGCTTGGCAAGACGCTTAAAGCT 1351
Db 1261 TCTGCAAGGTGACCCCGGCTGAGATACCTTCTTCCGCTTGGCAAGACGCTTAAAGCT 1320
QY 1352 TCCGAAAGGCAATTTCTCAAGAAAGTCTCTACTCTGTACGAGGAGGAGGACAGAGGAGCC 1411
Db 1321 TCCGAAAGGCAATTTCTCAAGAAAGTCTCTCTACTCTGTACGAGGAGGAGGAGGAGGAGCC 1380
QY 1412 TCAAGAACTGACACCGGTGCTGATCTTCCAGGAGACCTGTCCCTGTATCCCTGCT 1471
Db 1381 TCAAGAACTGACACCGGTGCTGATCTTCCAGGAGACCTGTCCCTGTATCCCTGCT 1440
QY 1472 CCCAAGGAAAGGCACTCTTTAAAGATCTCGACAGCTGATCTGTGTTACTACTCTCTCC 1531
Db 1441 CCCAAGGAAAGGCACTCTTTAAAGATCTCGACAGCTGATCTGTGTTACTACTCTCTCC 1500
QY 1532 AGAGCCCAAGGAGTCTGGGGAACCTTGAAGCCAGGATGTGTTTGTGATGGGAGCC 1591

Db 1501 AGAGCCCAAGGAGTCTGGGGAACCTTTAGACGCCAGTATGTGTTGTAGTGGGAGCC 1560
QY 1592 CGTGAAGCAGAACTTCCACAGGCTTCAAGGCTCTCTCTCCACCGCAAGGCAATTTCAA 1651
Db 1561 CGTGAAGCAGAACTTCCACAGGCTTCAAGGCTCTCTCTCCACCGCAAGGCAATTTCAA 1620
QY 1652 ACTCAATGGAATTTCTCCCGGACAGCTTTAGAAAGGACCTAACCCCTAGACCTTTGGCTC 1711
Db 1621 ACTCAATGGAATTTCTCCCGGACAGCTTTAGAAAGGACCTAACCCCTAGACCTTTGGCTC 1680
QY 1712 CTTGAGCAACATGAGCTCTCCATCCCTGATCGAGCCGAGCGCCCTCAAGGAGCTGT 1771
Db 1681 CTTGAGCAACATGAGCTCTCCATCCCTGATCGAGCCGAGCGCCCTCAAGGAGCTGT 1740
QY 1772 GAGTGAAGCAGACATCTGTCTCTCGAGTCTTTGACCAATTTGACTGTGCTGACGCT 1831
Db 1741 GAGTGAAGCAGACATCTGTCTCTCGAGTCTTTGACCAATTTGACTGTGCTGACGCT 1800
QY 1832 TCCCGAAACCCCACTGAGGAGCTGTGTCTGTGTGACCAACTGAGGAGGCTTGAAGCGCC 1891
Db 1801 TCCCGAAACCCCACTGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
QY 1892 TCCCTCAAGAGTCTGAAAGCGATGTGTGAGAGAAATCCTTGGGGAATGTGCTGCTTTCT 1951
Db 1861 TCCCTCAAGAGTCTGAAAGCGATGTGTGAGAGAAATCCTTGGGGAATGTGCTGCTTTCT 1920
QY 1952 GACAGACTGCAAGAGGTGACTGACAGCTTACAGCAAGCCCTTGAATCTGCTCAAGCT 2011
Db 1921 GACAGACTGCAAGAGGTGACTGACAGCTTACAGCAAGCCCTTGAATCTGCTCAAGCT 1980
QY 2012 CAGCTGAGGAAGGAGATGTGTGCTGTATGTGGGATAGGCTGTGAGAGGCTTGGCAAGG 2071
Db 1981 CAGCTGAGGAAGGAGATGTGTGCTGTATGTGGGATAGGCTGTGAGAGGCTTGGCAAGG 2040
QY 2072 AACCTGAGGTGCGATTTCTCTCAAGTATAGTATCAAGAGGCTCTACCTGACGCT 2131
Db 2041 AACCTGAGGTGCGATTTCTCTCAAGTATAGTATCAAGAGGCTCTACCTGACGCT 2100
QY 2132 GACTGAACTTGAAGATGAGAGAAATGCAATTTGATGTGAAAGAAATGGAAACCTTTGCT 2191
Db 2101 GACTGAACTTGAAGATGAGAGAAATGCAATTTGATGTGAAAGAAATGGAAACCTTTGCT 2160
QY 2192 GCCCGAGTGTATATGTGGGTGAGCTGAAAGTGTCTACCTCTTTGTGCAATGAGTCA 2251
Db 2161 GCCCGAGTGTATATGTGGGTGAGCTGAAAGTGTCTACCTCTTTGTGCAATGAGTCA 2220
QY 2252 CCCATGACATTTCCCAACCCCTGTTCTCTGAGTGAACCTTGCATATAGTTTCTGTTCCATC 2311
Db 2221 CCCATGACATTTCCCAACCCCTGTTCTCTGAGTGAACCTTGCATATAGTTTCTGTTCCATC 2280
QY 2312 AACCAACGAGGTAGAACCTCTGACTTCTGAGAGGTAAATGTGTATGTGACATGCAATATTT 2371
Db 2281 AACCAACGAGGTAGAACCTCTGACTTCTGAGAGGTAAATGTGTATGTGACATGCAATATTT 2340
QY 2372 AGAGAGGAACAGCTCTGTGTTCCATCTGTGCTGTGTGCAATCTCAAAACCTGGGAG 2431
Db 2341 AGAGAGGAACAGCTCTGTGTTCCATCTGTGCTGTGTGCAATCTCAAAACCTGGGAG 2400
QY 2432 ACTCGACGCTGTTGACTTCAATCTCAAGGAGACAGATGAGCCCTGAGACCCCATCTTAG 2491
Db 2401 ACTCGACGCTGTTGACTTCAATCTCAAGGAGACAGATGAGCCCTGAGACCCCATCTTAG 2460
QY 2492 ATCTCAGAGACTTGAACCTTGAAGCTGTTCTTATGATCCAGATGTGATGATGCTGT 2551
Db 2461 ATCTCAGAGACTTGAACCTTGAAGCTGTTCTTATGATCCAGATGTGATGATGCTGT 2520
QY 2552 TTCTCAGGCGCAAGGAGCTTGAATGTGCTGACTTATTTATTTTGTGATTTCTCACTT 2611
Db 2521 TTCTCAGGCGCAAGGAGCTTGAATGTGCTGACTTATTTATTTTGTGATTTCTCACTT 2580
QY 2612 CTGTTTTTGGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGCTTTC 2671
Db 2581 CTGTTTTTGGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGCTTTC 2640

QY	2672	AAATATGGAATGCGTGTCTTGGGGAACTCAACTGTCACATGAAGTTATGTACAAG	2731
Db	2641	AATAATGGAATGCGTGTCTTGGGGAACTCAACTGTCACATGAAGTTATGTACAAG	2700
QY	2732	AAGATTTGGCAATGATGTCCTCTATTCAAAGGGGGTGGGGCGTTTTCAAATGATG	2791
Db	2701	AAGATTTGGCAATGATGTCCTCTATTCAAAGGGGGTGGGGCGTTTTCAAATGATG	2760
QY	2792	TCTTGAGCACTGTCGTGAATTGAAGTCTCAATCCCTTCAACCCAAAGCTGGCCACCTCC	2851
Db	2761	TCTTGAGCACTGTCGTGAATTGAAGTCTCAATCCCTTCAACCCAAAGCTGGCCACCTCC	2820
QY	2852	CTCATCTTCATCTGTGGCCAAAAA	2896
Db	2821	CTCATCTTCATCTGTGGCCAAAAA	2865
RESULT 3			
AAAD31710			
ID	AAAD31710	standard; cDNA; 2929 BP.	
XX	AAAD31710;		
AC			
XX			
DT	18-JUN-2002	(first entry)	
XX			
DE	Rat SNF1/AMPK-Related Kinase (SNARK) cDNA.		
XX			
RW	Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;		
KM	glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;		
XX	hyperglycaemic; drug screening; hypoglycaemia; ss.		
XX			
OS	Rattus sp.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	83..1975	
FT		/*tag= a	
FT		/product= "Rat SNARK protein"	
XX			
FN	W0200212456-A2.		
XX			
PD	14-FEB-2002.		
XX			
PF	02-AUG-2001; 2001WO-CA001109.		
XX			
PR	03-AUG-2000; 2000US-0222650P.		
PR	12-MAR-2001; 2001US-0274613P.		
PR	28-MAR-2001; 2001CA-02340783.		
XX			
PA	(ONEO-) 1149336 ONTARIO INC.		
XX			
F1	Drucker DJ, Rosen CF, Lefebvre DL;		
XX			
DR	WPI; 2002-241747/29.		
XX			
DR	P-PSDB; AAE19885.		
XX			
PT	AMK (AMP-activated protein kinase)-related kinase, designated SNARK		
PT	polypeptides and polynucleotides, useful for treating or preventing		
PT	diabetes, or other disorders of lipoprotein production leading to		
PT	increased levels of cholesterol.		
XX			
PS	Example; Fig 2; 94pp: English.		
XX			
CC	The invention relates to an AMPK (AMP-activated protein kinase)-related		
CC	kinase, designated SNARK polypeptides and polynucleotides. SNARK		
CC	(SNIF/AMP-activated protein kinase) is involved in stress response to		
CC	glucose deprivation. The polynucleotides are useful for expressing SNARK		
CC	protein in isolated form or as a protein conjugate. Activation of SNARK		
CC	stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and		
CC	in other cell types such as heart and skeletal muscles, as well as		
CC	increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted		
CC	to have insulin-like effects that would enhance the disposal of glucose		
CC	into muscle and reduce plasma glucose for the treatment of diabetes and		

Db 849 GGCAGATCATTAACCTGCTGTAACAAATCATAGCGGGCTTACCGAGAGCCGTGCA 908
QY 949 AGCGCTCCAGATGCTGTGGCTGTATCCGGTGGCTGTATATGTGAACCCACCCGTGCG 1008
Db 909 AACCGTGTAGTGCCTGTGGCTGTATCCGGTGGCTGTATATGTGAATCCATCCGTGCG 968
QY 1009 CCAACAGTGAAGATGTAGCAGATGATGTGTGTGCACTGGGGTGTACACCAACCGAGATCG 1068
Db 969 CCACTGTGAGAGATGTAGCAGATGATGTGTGTGCACTGGGGTGTACACCAACCGAGATGTG 1028
QY 1069 GGGAAACAGAAAGCCCTGCTGAGGGGTGGGCAACCTTAGTGTGACTTTTGCCGGGCTTCA 1128
Db 1029 GGGAAACAGAAAGCTCTGAGAGGGGTGGGCAACCTTAGGGTGTGACTTTTGCCGGGCTTCA 1088
QY 1129 TGGGGAAGTGTATGCTGCTCTCGGCCCTCTCTGAGAGATGAGCCAAAGTGTGCA 1188
Db 1089 TGGGGAAGTGTATGCTGCTCTCGGCCCTCTCTGAGAGATGAGCCAAAGTGTGTA 1148
QY 1189 GCTTCTTCAAGCAGCAGTGCCTGGGAGTGTGAGACATGTATCTGGGCTGTGAGCGGCAAC 1248
Db 1149 GCTTCTTCAAGCAGCAGTGTGCCTGGGAGTGTGAGACAGCGGAGCTGTGAGCGGCAAC 1208
QY 1249 ATTCTCTTAAGAAATCCGAAAGAGATGACATGTGCTCAAAATCTGCAAGGTGACCCGG 1308
Db 1209 ATTCTCTTAAGAAATCCCGCAAGAGATGACATGTGCTCAAGATCTGTGCAAGATGACCCAG 1268
QY 1309 CTGAGATACCTCTTCTGCTGCTGTGCAAGAGACCTTAAAGCTTCCGAAAGGCAATTCTCA 1368
Db 1269 TTTAAGATATCTTCTCTGCTGTGCAAGAACAGCTTCAAGCTTCCGAAAGGTATCTCTCA 1328
QY 1369 AGAAAAAGTCTCTAATCTGCTGAGGGAGTGTACAGAGAGACCTTCAAGAACTCAAGCCG 1428
Db 1329 AGAAAAAGTCTCTCTCTCAATCGGGAGGTGTACAGAGAGGCTTCAAGAACTCAAGCCAG 1388
QY 1429 TGCCTGATATCTCAGGAGCAGCTGTCTCCCTGCTGTATCCCTGCTCCAAAGGAAAGCATCC 1488
Db 1389 TGTCTCAATATCCCAAGGAGCAGCTGTCTCTCTGTATACCTCTGCTCCAAAGGAGGAGATTC 1448
QY 1489 TTAAGAAATCTCAGCAGCGTGAATCTGTATCTAATCTCTCTCAAGACCCAGCAGATCTG 1548
Db 1449 TTAAGAAATCTCGGCAAGCGTGAATCTGTATCTAATCTCTCTCAAGACCCAGTGAATCTG 1508
QY 1549 GGGAACTCTTAAGAGCCAGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1608
Db 1509 GGGAACTCTTAAGAGCCAGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1568
QY 1609 CACAGGCTTCAAGGAGCTCTCTCTCCAGCAGGAGCATTTCTCAAACTGAATGGCAAGTCT 1668
Db 1569 CACAGGCTTCAAGGAG--CGCTCTCATGTGCAAGGAGCATCTCAAACTGAATGGCAAGTCTT 1625
QY 1669 CCGGCAAGCTTAAAGAGGACATACCTCTAGCATTGTGGCTCCCTGAGCAACATGGCT 1728
Db 1626 CCGGCAAGCTTAAAGAGGACATGCTCCCTGAGCATCTTTGGCTCCCTGAGCAACATGGCT 1685
QY 1729 CTTCCCATCTGTGAGCCGCGCCAGCCGCTTGAAGGGCTGTGATGTGAGAGCAGATCC 1788
Db 1686 CCGCTCATCTTAAGCCCGGAGCAGCGCTCTCTGAGAGCTGTGATGTGAGAGCAGATCC 1745
QY 1789 TGTCTCTCCAGTCTTGAACCAATTTGACATTTGCTGTGAAAGTCTTCCGAAACCCCACTGA 1848
Db 1746 TGTCTCTCCAGTCTTGTGACCAATTTGACATTTGCTGTGAAAGTCTTCCGAAACCCCACTGA 1805
QY 1849 GGGGCTGT 1908
Db 1806 GGAAGCTGT 1865
QY 1909 AGCGATGT 1968
Db 1866 AAGCATGT 1925
QY 1969 TGAATCAAGCTTCAAGCAGCCTTGAAGATGTGTCTCAAAAGCTGTGAGGAAAGGAGAGA 2028

Db 1926 TGAACAGAGCCTTACAGCAAGCCCTAGAAATCTGCTGAAAGCTGACGTGAGAGAGAGG 1985
QY 2029 TGTGCTCTAGT-ATGGGGTGAAGGCTGTGAAGAGGTTTGCAGAGAACCTTGGGTGAGTT 2087
Db 1986 CAGTCCCCAGTGTAGGGTGTAGACTCTTGAAGGGGTTTGCAGAGAACCTTGGGT-AGATT 2044
QY 2088 CTTCCAGTGAATGTAGTACATCAAG----GGCTTACGTGTGAGCCCTGACTGAACCTGA 2143
Db 2045 CCGGAGGTTGTAGATGATCAAGAACTCTCTCTGTCTTCAAGCTTGAATGAACCTGG 2104
QY 2144 AAGATGAGAGAAATCGCATGTATGTGAGAAAGAAATGGAACCTTGTGCTGCCGAGTGTGA 2203
Db 2105 AGCGTGAAGAAATAGAGAGATGTGAAAGAGCTGACCTTACAGAGATCTGACATGATG 2154
QY 2204 TAGTGGGT--GGCTGAAGGTGCTTACCTCTTGTGCAATGATGTGACCCATGACA 2260
Db 2165 TGAAGGCAAGAGACTGAAGGTGCTTACCTCTTATGCT--TGAAGTCAACCATGGCA 2222
QY 2261 TTTCCCAACCTGTCTCTGCTGACAC-----TTCAATTAAGTTCTGTTCATCA 2312
Db 2223 TCTCCC-CCCTGCTCTGCAAGTGTGAGGGTGTACCAATTAAGTTCTGTTCGACATCG 2281
QY 2313 ACCACAGGGTTGAACCTTGAATCTCTGGAGGTAATGTGTGTGATGCTGCAATTTA 2372
Db 2282 ACCACAGGGTTGAACCTTGAATCTCTGGAGGTAATGTGTGTGATGCTGCAATTTA 2341
QY 2373 GAGAGAAACGCTCTGTGTTTCATCTGTGCTGTGATCTCAAAAGACTGGGAGAGA 2432
Db 2342 AAGAGAAACAGCTCTGTGTTTCATCTGTGCTGTGATCTCAAAAGACTGGAGAGA 2401
QY 2433 CTGGAAGCCTGTGTGATCTCATCTCAAGGGAGACAGATCCCTGGAACCCCATCTTGA 2492
Db 2402 CTC-AATGCTGTGTTCACTTCATCTCAAGGGAGC----- 2434
QY 2493 TCTCAGAGACTTGAACCTTGAAGCTGTCTTGAACCAAGATGTGATGAT-GCCTGT 2551
Db 2435 -CTCAGAGACTTGAACCTTGAAGCTGTCTTGAACCAAGATGTGATGATGATGCTGT 2493
QY 2552 TTCTCAGGCCAAGGAGCCTTGAATGTGCTGACTTATTTATTTTGTGTGATCTCACTT 2611
Db 2494 TTCTCAGGCCAAGGAGCCTGAATGTGCTGACTTATTTA--TTTGTGATCTCACTT 2551
QY 2612 CTGTTTTTGTGTT----- 2624
Db 2552 CTGTTTTTGTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2611
QY 2625 -----TTGTTTTTGTGTTTGTGTTTTTGAAGATTTTGTGCTTCAATATG 2678
Db 2612 TGTGTTTTGTGTTTTGTGTTTTGTGTTTTTGAAGATTTTGTGCTTGTGATATG 2671
QY 2679 TGAATGCTGTGTTCTGGGAACTCACTGTGCACTGAAGTTATGTACAGAGAAATTT 2738
Db 2672 TGAATGCTGTGTTCTGGGAAAGCACTGTGATGAAATGTGTGTGTGTGTGTGTGTGT 2731
QY 2739 TGGCAATGATGCTCTTATTAAGGGGGGTGGGGGCTTTTCAAAATGTATGTCTTGAG 2798
Db 2732 TGGCAGATGATCTTCTTAAT-----GGGGGGTGGCTTTTCAAGATGTATGCTTGAG 2783
QY 2799 CACTGCTGTGATGTAGTCTCAATCCCTTACACCAAGGCTGGCAACCTCTCATCT 2858
Db 2784 CACTGCTGTGATGTGCTGTCTGTCTCTCAACCAAGGCTGTGCAACCTCTCATCT 2843

RESULT 4
ABD33082
ID ABD33082 standard, DNA, 37278 BP.

XX ABD33082;
AC
XX
XX
DT 18-NOV-2004 (first entry)
XX
XX
DE Murine cancer-associated (CA) gene MD07-004.

KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KM de; cancer; cytostatic.
XX Mus musculus.
OS
PN W02004058146-A2.
XX
PD 15-JUL-2004.
XX
PF 15-DEC-2003; 2003WO-US040081.
XX
PR 17-DEC-2002; 2002US-00322281.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
PI Morrie DW, Malandro MS;
XX WPI; 2004-499109/47.
XX
PT Novel human cancer associated protein encoded within open reading frame
XX of cancer associated gene, useful as targets for diagnosing cancer.
PS Disclosure; SEQ ID NO 21; 182pp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 37278 BP; 8909 A; 9465 C; 8785 G; 9687 T; 0 U; 432 Other;
Query Match 66.8%; Score 1938.8; DB 13; Length 37278;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 955 CCGATGCTGTGCTGCTGATCCGATGCTGTTAATGTGAACCCACCCGTCGGCCACAC 1014
DB 25129 CAGATGCTGTGCTGCTGATCCGATGCTGTTAATGTGAACCCACCCGTCGGCCACAC 25188
QY 1015 TGGAGATGTATGCCAGTCACTTGTGGGTCAACTGGGGTTTACACACCGGAGTCGGGAAC 1074
DB 25189 TGGAGATGTATGCCAGTCACTTGTGGGTCAACTGGGGTTTACACACCGGAGTCGGGAAC 25248
QY 1075 AGGAAGCCCTGCGGAGGAGTGGGACACCTTAATGTGAACCTTGGCCGGGCTCCACATGGGCG 1134
DB 25249 AGGAAGCCCTGCGGAGGAGTGGGACACCTTAATGTGAACCTTGGCCGGGCTCCACATGGGCG 25308
QY 1135 ACTGTTAAGTGTGCTCTCTGCGCCCTCTCTGAGAAATGAGACCAAGGTGTGACGCTTCT 1194
DB 25309 ACTGTTAAGTGTGCTCTCTGCGCCCTCTCTGAGAAATGAGACCAAGGTGTGACGCTTCT 25368
QY 1195 TCAAGCAGCAGTGTGCGGAGGTGAAGACATGTAACCTGGGTGAGCGGCAACATTTCTC 1254
DB 25369 TCAAGCAGCAGTGTGCGGAGGTGAAGACATGTAACCTGGGTGAGCGGCAACATTTCTC 25428
QY 1255 TTAGAAGTCCCGAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCGGCTGAGG 1314
DB 25429 TTAGAAGTCCCGAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCGGCTGAGG 25488
QY 1315 ATACCTCTTCTGCGCTGCGCAAGACGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAA 1374

DB 25489 ATACCTCTTCTGCGCTGCGCAAGACGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAA 25548
QY 1375 AGTCTCTACCTCTGTCAGGAGAGGTATACAGAGACCCCTCAGAACTCAGACCGGTGCTG 1434
DB 25549 AGTCTCTACCTCTGTCAGGAGAGGTATACAGAGACCCCTCAGAACTCAGACCGGTGCTG 25608
QY 1435 ATACTCCAGGAGCCTGTGCTGCTGTATCTCTGCTCTCCAGGAAAGGCAATCTTAAAGA 1494
DB 25609 ATACTCCAGGAGCCTGTGCTGCTGTATCTCTGCTCTCCAGGAAAGGCAATCTTAAAGA 25668
QY 1495 AGTCTCCAGGAGCCTGTGCTGCTGTATCTCTGCTCTCCAGGAGCCTGCTGGGAGAC 1554
DB 25669 AGTCTCCAGGAGCCTGTGCTGCTGTATCTCTGCTCTCCAGGAGCCTGCTGGGAGAC 25728
QY 1555 TCTTAGACGCGCAGTGTGTGTGTGATGGGAGACCCGCTGAGACAGAAAGTCTCCACAGG 1614
DB 25729 TCTTAGACGCGCAGTGTGTGTGTGATGGGAGACCCGCTGAGACAGAAAGTCTCCACAGG 25788
QY 1615 CTTGAGGAGCTCTCTCCACCGCAGGAGGCAATTTCAAACTCAATGACAAAGTTCTCCGCA 1674
DB 25789 CTTGAGGAGCTCTCTCCACCGCAGGAGGCAATTTCAAACTCAATGACAAAGTTCTCCGCA 25848
QY 1675 CAGCTTAAAGAGCACTACCTTAGACACTTTGGCTCCCTGAGACCAATGCGCTCTCTCC 1734
DB 25849 CAGCTTAAAGAGCACTACCTTAGACACTTTGGCTCCCTGAGACCAATGCGCTCTCTCC 25908
QY 1735 ATCTGAGGCGCGGCGCAGCGCCCTCAGGAGGCTGTGAGTGAAGAGACAGCATCTGCTCT 1794
DB 25909 ATCTGAGGCGCGGCGCAGCGCCCTCAGGAGGCTGTGAGTGAAGAGACAGCATCTGCTCT 25968
QY 1795 CCGAGTCTTTGACCAATTTGACCTTGTGCTGACACTTCCGAAACCCCACTGAGGAGCT 1854
DB 25969 CCGAGTCTTTGACCAATTTGACCTTGTGCTGACACTTCCGAAACCCCACTGAGGAGCT 26028
QY 1855 GTGTGTCTGTGACCACTGAGGAGGCTTTGAGACGCTCTCTCAGAAAGTCTTAAAGCAT 1914
DB 26029 GTGTGTCTGTGACCACTGAGGAGGCTTTGAGACGCTCTCTCAGAAAGTCTTAAAGCAT 26088
QY 1915 GTGTGAGGAATCTTGTGGGAGATGAGCTTTCTGTGACGACCTGCAAGAGGTGATG 1974
DB 26089 GTGTGAGGAATCTTGTGGGAGATGAGCTTTCTGTGACGACCTGCAAGAGGTGATG 26148
QY 1975 CAGCTTACAGCAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAGAGATGTGTC 2034
DB 26149 CAGCTTACAGCAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAGAGATGTGTC 26208
QY 2035 CTTAGTATGGGAGTGTCTTGAAGGCTTTGACAGAGAAACCTTGGGTGCAATTTCTCCAG 2094
DB 26209 CTTAGTATGGGAGTGTCTTGAAGGCTTTGACAGAGAAACCTTGGGTGCAATTTCTCCAG 26268
QY 2095 TGAATTAAGTATCAACAGGAGCTTACGCTGACAGCTGACCTGAACCTGAAGATGAGAGA 2154
DB 26269 TGAATTAAGTATCAACAGGAGCTTACGCTGACAGCTGACCTGAACCTGAAGATGAGAGA 26328
QY 2155 AATGCAATTGATGTGAAGAAAGGAATGGGAACCTTGTGCTGCGGAGTGTATATGTGGGCTG 2214
DB 26329 AATGCAATTGATGTGAAGAAAGGAATGGGAACCTTGTGCTGCGGAGTGTATATGTGGGCTG 26388
QY 2215 CTTGAGAGTGTCTACCTCTTGTGTCATGAGTGTCAACCAATGACATTTTCCACCTGTT 2274
DB 26389 CTTGAGAGTGTCTACCTCTTGTGTCATGAGTGTCAACCAATGACATTTTCCACCTGTT 26448
QY 2275 CTCTGCTGACCTTTCATTAAGTTTCTGTTTCCATCAACCAAGGTTTGAAGACCTTGA 2334
DB 26449 CTCTGCTGACCTTTCATTAAGTTTCTGTTTCCATCAACCAAGGTTTGAAGACCTTGA 26508
QY 2335 CTTCTGGAGGTATGTATGTAGTACCTGCAATTTTGAAGAGAAACAGGCTCTGTTT 2394
DB 26509 CTTCTGGAGGTATGTATGTAGTACCTGCAATTTTGAAGAGAAACAGGCTCTGTTT 26568
QY 2395 CCATCTGTGCTGTGCTGCTCAAAAGCTTGGAGAGCTGGAACCGCTGTTTGAATTTCA 2454
DB 26569 CCATCTGTGCTGTGCTGCTCAAAAGCTTGGAGAGCTGGAACCGCTGTTTGAATTTCA 26628

QY 2455 TCTCAAGGAGCAGATGCCCCCTGGACCCCACTTAAATCTCAGAGACTTGACCTTCAA 2514
DB 26629 TCTCAAGGGAGCAGATGCCCCCTGGACCCCACTTAAATCTCAGAGACTTGACCTTCAA 26688
QY 2515 GCTGTTCTTGAATGACCCAGATGATGATGCTCTGTTTCTCAGGCCAAGGAGCTTGA 2574
DB 26689 GCTGTTCTTGAATGACCCAGATGATGATGCTCTGTTTCTCAGGCCAAGGAGCTTGA 26748
QY 2575 ATGTGCTGACTTAAATTTATTTTGTGATGATCTGATCTGTTTGTGTTTGTGTTT 2634
DB 26749 ATGTGCTGACTTAAATTTATTTTGTGATGATCTGATCTGTTTGTGTTTGTGTTT 26808
QY 2635 TGTGTTGTTTGTGTTTAAAGAAATTTGCTGCTTTCATATATGGAATGCTGTTCTG 2694
DB 26809 TGTGTTGTTTGTGTTTAAAGAAATTTGCTGCTTTCATATATGGAATGCTGTTCTG 26868
QY 2695 GGGAACTCACTGTCGCACTGAAGTTATGTAACAGAGAAATTTGGCAATGATGCTCCT 2754
DB 26869 GGGAACTCACTGTCGCACTGAAGTTATGTAACAGAGAAATTTGGCAATGATGCTCCT 26928
QY 2755 CTATTCAGAGGGGGGTGGGGGGGTTTTCMAATGTAATGTTGAGCACTGTCGATTGAG 2814
DB 26929 CTATTCAGAGGGGGGTGGGGGGGTTTTCMAATGTAATGTTGAGCACTGTCGATTGAG 26988
QY 2815 TCTCCAGTCCCTTTCACACCCCAAGGCTGGCCACCTCCTCATCTTTCATCTGTGCCAAA 2874
DB 26989 TCTCCAGTCCCTTTCACACCCCAAGGCTGGCCACCTCCTCATCTTTCATCTGTGCCAAA 27048
QY 2875 AAAAAAAAAAAAAAAAAAAAAA 2896
DB 27049 AAAAAAAAAAAAAAAAAAAAAA 27070

RESULT 5
ID ABRK14000 standard; cDNA; 3353 BP.
AC ABRK14000;
XX 02-JUL-2002 (first entry)
DE cDNA encoding human protein kinase 3700.
KW Human, protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
KW cell proliferation disorder; cell differentiation disorder; carcinoma;
KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;
KW cytostatic; antiatherosclerotic; gene; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 157..2043
FT CDS /*tag= a
FT /product= "Protein kinase 3700"
FT /note= "The coding region (not including the terminator
FT codon) is specifically claimed in claims 7 and 28"
XX
XX WO200224921-A2.
XX
XX 28-MAR-2002.
XX 25-SEP-2001; 2001WO-US030115.
XX 25-SEP-2000; 2000US-0234922P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Curtis RAJ, Galvin KM;
XX PI
XX WPI; 2002-352007/38.

DR P-RSDB; AAU79652.
XX
XX Use of modulators of activity of 3700 protein for making medicament for
PT e.g., modulating protein phosphorylation or cell signalling, or for
PT treating or preventing cellular proliferative and/or differentiative
PT disorders.
XX
XX Claim 28; Fig 1; 115pp; English.
XX
XX The present invention relates to the isolation of a novel human protein
CC kinase designated 3700, and the polynucleotide sequence encoding it. The
CC invention also describes the use of a modulator of the activity of
CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
CC composition for modulating the ability of a cell to phosphorylate an
CC amino acid residue of a substrate protein. Modulators of protein kinase
CC 3700 activity are useful for modulating protein phosphorylation, cell
CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
CC angiogenesis, tissue repair, tissue regeneration, establishment or
CC progression of atherosclerosis, and signalling across the blood-brain
CC barrier. The polynucleotide and polypeptide molecules for protein kinase
CC 3700 may be used as diagnostic targets and therapeutic agents for
CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
CC curing PK-related disorders and cellular proliferative and/or
CC differentiative disorders (e.g. haematopoietic neoplastic disorders,
CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
CC polynucleotide sequence can be used to express protein kinase 3700, to
CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for
CC tissue typing, in forensic biology, and as surrogate markers. The present
CC sequence encodes human protein kinase 3700
XX
SQ Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;
Query Match 47.5%; Score 1379; DB 6; Length 3353;
Best Local Similarity 81.8%; Pred. No. 9.3e-295;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;
QY 82 CTACTGCTGCCCAACCACTTCACCTGCGGTCCCGCACCAATGAGTCCGTGCTTAC 141
DB 116 CTACTGATTCCTCCCTGCGCCCTTGTCTCCTCTGCTGCGCATGAGATCGCTGTTTTCG 175
QY 142 TCCAGCGCCCGAGCGAGGCTCCCTCGGCTCCCGCCCTGGAGAGAGCGCCGCGCGC 201
DB 176 CGCGGCGCTCGGCGCCCACTTCCTCGGCGGAGCTTA-----GCCCGGCGC 223
QY 202 TGGCGGAGCGGCTCATCAAGTCCCTTAACTCTGATGAGAGAGCAGCGGTGAAGCGC 261
DB 224 TGGGGAAGGGCTGATCAAGTCGCGCAAGCCCTTAAGAGAGAGCGGTGAAGCGC 283
QY 262 ACCATCAAAACACACACTCTGGGACCGCTACGAGTTCTTGAGAGCGCTGGGCAAGGCA 321
DB 284 ACCACCAAAAGCACACCTCGGACCGCTACGAGTTCTTGAGAGCGCTGGGCAAGGCA 343
QY 322 CCTACGGGAAGGTGAAGAGGACGAGAGCTCGGGGCGTCTGTGTCATCAAGTCCA 381
DB 344 CTTACGGGAAGGTGAAGAGGCGGAGAGCTCGGGGCGTCTGTGTCATCAAGTCCA 403
QY 382 TCAAGAAAGCAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGGAGATTGAGA 441
DB 404 TCCGGAAGGCAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGGAGATTGAGA 463
QY 442 TCATGCTTCACTCAACCACTCCCACTCATCTGCGATTCATGAAGTGTGAGAAATGCA 501
DB 464 TCATGCTATCACTCAACCACTCCCACTCATCTGCGATTCATGAAGTGTGAGAAAGCA 523
QY 502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
DB 524 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 583
QY 562 AGCGGCAAGCGCTGAGTGAAGCGGAGCGGAGGATTTCTTCGACAGATGATGATGATGAT 621
DB 584 AGCGGCAAGCGCTGAGTGAAGCGGAGGAGGATTTCTTCGAGATGATGATGATGATGAT 643
QY 622 TGCACTAATGCGCACCAAGACGGAATGCTTCAACGAGATTCGAAGCTGAAAAACATCTTC 681

Db	1715	CCTTCGAGCTCCCTGGATGAAGTACGTGCAGCCCACTGCAGCCCTCGAGCCCGAGCCGACCT	1774
Qy	1762	CAGGGGCTGTGAGTGAAGACAGCATCTGTCTCTCCAGTCTCTTTGACCAATTGGACTTGC	1821
Db	1775	CAGGGGCTGTGAGTGAAGACAGCATCTGTCTCTCTGAGTCTTTGACCACTGACTTGC	1834
Qy	1822	CTGAACGCTTCCCGAAACCCCACTGAAGGGGCTGTGCTGTGAGCAACACTGAGGGGCG	1881
Db	1835	CTGAACGCTTCCCGAGAGCCCACTGAGGGGCTGTGCTGTGAGCAACACTTCAAGGGCG	1894
Qy	1882	TTGAGCAGCCTCCCTCAGAG-----GTCTGAAGCGATGTGTGACAGATCT	1929
Db	1895	TTGAGAGAGCCCCCTCAGAGGGGCGCTGGAAGCTGCTCAGAGGCGTGGCGGACGATCTT	1954
Db	1955	TGGGGGACAGCTGCTTTTCCCTGACATGCGCAGAGGTGACAGACACTTACCGACAGG	2014
Qy	1990	CCCTTAGAATCTGCTCAAGTCAAGCTGAGAGGAGAGATGTGCGCTTATATGGGTAG	2049
Db	2015	CACTAGAGGCTGTGCTCAAGTCACTGAGTGAAGTGAAGCATGTGCCCCAG-CCGGTCAAG	2073
Qy	2050	GCTCTGAGAGGATTTCAGAGAGAACCTGGGTGGATTTCTCC	2092
Db	2074	GCTCTCAGATCAGCTGTGTGCACCCCGAGGGAGATGCTTC	2116
RESULT 6			
ID	ADL14160		
AC	ADL14160	standard, cDNA, 3353 BP.	
XX	ADL14160;		
XX	17-JUN-2004	(first entry)	
DE	Novel human gene 3700 cDNA.		
KW	cytostatic; cardiant; hypotensive; antianginal; osteopahic;		
KW	antiarthritic; antithematic; neuroprotective; antiinflammatory;		
KW	angioprotect; antiaesthetic; cardiovascular; vitricide; analgesic; CNS;		
KW	angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;		
KW	neurotropic; antithyroid; dermatological; immunomodulator;		
KW	cell proliferation disorder; cell differentiation disorder;		
KW	brain disorder; platelet disorder; breast disorder; colon disorder;		
KW	kidney disorder; renal disorder; lung disorder; ovarian disorder;		
KW	prostate disorder; cervical disorder; spleen disorder; thymus disorder;		
KW	thyroid disorder; testes disorder; haematopoietic disorder;		
KW	pancreatic disorder; skeletal muscle disorder; skin disorder;		
KW	dermal disorder; bone metabolism disorder; immune disorder;		
KW	inflammatory disorder; cardiovascular disorder;		
KW	endothelial cell disorder; liver disorder; viral disease; pain disorder;		
KW	metabolic disorder; neurological disorder;		
KW	central nervous system disorder; erythroid disorder;		
KW	blood vessel disorder; angiogenic disorder; cancer; heart failure;		
KW	hypertension; angina; osteoarthritis; rheumatoid arthritis;		
KW	multiple sclerosis; Crohn's disease; psoriasis; asthma;		
KW	cell proliferation; cell differentiation; cell growth; cell division;		
KW	human; gene; ss.		
OS	Homo sapiens.		
XX			
XX			
PN	US2004058355-A1.		
XX			
PD	25-MAR-2004.		
XX			
PF	25-APR-2003; 2003US-00423543.		
XX			
PR	30-SEP-1998; 98US-00163821.		
PR	27-JAN-1999; 99US-0117580P.		
PR	25-MAR-1999; 99US-00276400.		
PR	30-JUN-1999; 99US-00365162.		
PR	09-SEP-1999; 99US-00392189.		
PR	05-OCT-1999; 99US-00412210.		


```

Db      1004 TGTGATGATGATACCCCAACCCGCGGCGCACTGTGAGATGTGAGCCAGTCATGTGTGG 1063
Qy      1042 TCAACTGGGGTTACACCAACCGAGTTCGGGGAAACAGAAAGCCCTGCTGAGGGGTGGCAAC 1101
Db      1064 TCAACTGGGGTTACACCAACCGAGTTCGGGGAAACAGAAAGCCCTGCTGAGGGGTGGCAAC 1123
Qy      1102 CTAGTGTGATTTTGGCCGGGCTTCATGCGGAGCTGTGTAAGTGTCTCTCGGCCCCC 1161
Db      1124 CTGGCAGTGAATCTGCGCGCGCTTCATGCTGACTGGCTCGGCGTTCTCCGCCCCC 1183
Qy      1162 TCCCTGAGAAATGAGCCCAAGGTGTGCACTTCTTCAAGCAGACAGTGTGCGGAGAGTGA 1221
Db      1184 TCCCTGAGAAATGAGCCCAAGGTGTGCACTTCTTCAAGCAGACAGTGTGCGGAGAGTGA 1243
Qy      1222 GCACTGATACCTGGGCTGAGAGCGGCAACATTTCTTAAAGAAATGCCGAAGAGATGACA 1281
Db      1244 GCACCAACCTCTGGGCTTGAAGCGGCAAGTTCCTTCAAGAAATCCCGCAGAGAAATGACA 1303
Qy      1282 TGGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTTCTTCTGCGCCCTGGCAAGACA 1341
Db      1304 TGGCCAGTCTCTCAACAGTGAACGCGGTGATGACATGCCCATGCGCTTGGCAAGACA 1363
Qy      1342 GCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCTTAACTCTGTCAGGGAGGTAC 1401
Db      1364 ACCTCAAGCTGCGCAAGGGCAATTTCTCAAGAAAGGTGTGCTGCTGCAAGAAAGGGTAC 1423
Qy      1402 AGGAGACCTCTCAAGAACTGACAGCGGTCCTGATACCTCAAGGGAGCGTGTCTGCTG 1461
Db      1424 AGGAGACCTCTCGGAGTCAAGCCCAATCCCTGGAGCCCAAGGGAGGCTGCTGCTGCTG 1477
Qy      1462 TATCCCTGCTCCCAAGAAAGGCAATCTTAAAGAAATCTGCAAGCGTGAATCTGCTTACT 1521
Db      1478 ---CCCTGCTCCCAAGAAAGGCAATTTCTCAAGAAAGCCCGCAGCGAGTCTGCTACT 1534
Qy      1522 ACTCTCTCCAGAGCCCAAGGCTGTGGGAACTCTTAAAGCGCAAGTGTGTTTGA 1581
Db      1535 ACTCTCTCCAGAGCCCAAGGCTGTGGGAACTCTTAAAGCGCAAGTGTGTTTGA 1594
Qy      1582 GTGGGAGCCCTGTGAGAGAGTCTCAAGGCTTCAAGGGCTCTCTCCACCGCAAG 1641
Db      1595 GTGGGAGTCCCAAGAGAGAGTCTCCGCAAGCTTCAAGGGCTCTCTCCATGCAAG 1654
Qy      1642 GCATTCTCAAACTCAATGCAAGTTCCTCCGACAGCCTTGAAGAGCACTACCTTAGCA 1701
Db      1655 GCATCTCAAACTCAATGCAAGTTCCTCCGACAGCCTTGAAGAGTGTGCGGCCCAACCA 1714
Qy      1702 CTTTGTGCTCTGAGCAACTGTGCTCTCCATCTGCAAGCCCGGCCAGCCGCTT 1761
Db      1715 CTTTGTGCTCTGAGCAACTGTGCTCTCCATCTGCAAGCCCGGCCAGCCGCTT 1774
Qy      1762 CAGGGGCTGTGAGAGAGCAAGCATCTGTCTCTCGAGTCTTGTGAACCAATTGACCTTGC 1821
Db      1775 CAGGGGCTGTGAGAGAGCAAGCATCTGTCTCTCGAGTCTTGTGAACCAAGCTTGC 1834
Qy      1822 CTGAACGCTTCCCGAAACCCCACTGAGGGCTGTGCTGTGAGAAACAACCTGAGGGGC 1881
Db      1835 CTGAACGCTTCCCGAAACCCCACTGAGGGCTGTGCTGTGAGAAACAACCTGAGGGGC 1894
Qy      1882 TTGAGAGCCTCTCTCAGAG-----GTCTGAAGCAATGTGTGAGAGAACTCT 1929
Db      1895 TTGAGAGCCTCTCTCAGAGGGCCCTGGAAGCTGTGAGGCGGTGGGGGAGATCTT 1954
Qy      1930 TGGGGGATAGTCTCTTTCTTGTACAGACTGCGCAAGAGTGACTGCAAGCAAG 1989
Db      1955 TGGGGGATAGTCTCTTTCTTGTACAGACTGCGCAAGAGTGACTGCAAGCAAG 2014
Qy      1990 CCTAAGAAATCTGCTCAAAAGCTCAAGTGAAGAAAGGAGATGTGCTGATGTGGGGTGA 2049
Db      2015 CACTGAGAGTCTGCTCAAAAGCTCAAGTGAAGAAAGGAGATGTGCTGATGTGGGGTGA 2073
Qy      2050 GCTGTGAGAGGTTTGCAGAGAAACCTGTGGTGGATTCCTC 2092
Db      2074 GCTCTCAGATGCAAGTGTGTGACACCCGAGGGGAGATGCTTTC 2116

```

```

RESULT 7
ID AAK94280
AAK94280 standard; cDNA; 3395 BP.
XX
AC AAK94280;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2918.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
EN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
DR P-PSDB; AAM93360.
XX
FT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2918; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
XX
SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
XX
Query Match 47.4%; Score 1375.8; DB 4; Length 3395;
Best Local Similarity 81.7%; Pred. No. 476-294;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
Qy 82 CTACTGTGCGCGCAACCACTTCACCTGCGGTCCTCCGCAAGTGAAGTGTGCTTAC 141
Db 77 CTACTGATTCCTCCGCGCGCTTCTCAACCTGCTGCTGCGCAGTGTGCTTTCG 136
Qy 142 TCCAGCGCCGAGCGCAGGCTTCTTCTGCGCTTCTGCGCTGCGGAGAGCGCCGCGCG 201
Db 137 CGCGGCTCTCGGCGCCCACTTCTTCTGCGCGGAGCTA-----GCCGCGCGC 184
Qy 202 TGGGGAAGGGCTCATCAAGTGTGCTTAACTCTGATGAAGAAAGAGGCGTGAAGCGC 261
Db 185 TGGGGAAGGGCTCATCAAGTGTGCTTAACTCTGATGAAGAAAGAGGCGTGAAGCGC 244
Qy 262 ACCATCAAAACACAACTGTGCGGCAACCGCTTGAAGTCTTGAAGAGCGTGTGGCAAGGCA 321
Db 245 ACCACCAAGAGCAAACTGTGCGGCAACCGCTTGAAGTCTTGAAGAGCGTGTGGCAAGGCA 304

```


QY	322	CTCAGCGGAAGGTGAAGAAGGACGAGAGAGCTCGGGGGGTCTGGTGGCCATCAAGTCA	381
Db	305	CTTACGGGAAGGTGAAGAAGGCGCGGGAGACTTCGGGGCGCTGGTGGCCATCAAGTCA	364
QY	382	TCAGGAAGAACAATAATCAAGATGAGCAGAGATCTGCTGCACTACGAGGAGATTGAGA	441
Db	365	TCCGGAGAAGCAAAATCAAAAGATGAGCAAGATCTGATGCACTACGAGGGAGATTGAGA	424
QY	442	TCATGTCCTCACTGAACCAACCCCACTCATTTGGCATCCATGAAATGTTTGAATAGCA	501
Db	425	TCATGTCATCACTCAACCACTCTCACTCATTTGGCATCCATGAAATGTTTGAACACCA	484
QY	502	GCAAGATTGTGATTTGTCAATGAGATATGCGCAGCGGAGATGTATGATTTACATCAGTG	561
Db	485	GCAAGATTGTGATTTGTCAATGAGATATGCGCAGCGGAGATGTATGATTTACATCAGTG	544
QY	562	AGCGGCCACGCGCTGAGTGAAGCGGGAGCGCCAGGCAATTTCTTCGACAGATCGTGTGCC	621
Db	545	AGCGGCCACGCGCTGAGTGAAGCGGGAGCGCCAGGCAATTTCTTCGACAGATCGTGTGCC	604
QY	622	TGCACTATGCGCACCAAGAACGGGATTCGTTCAACCGAGATCTCAAGCTGGAAAAATCTTTC	681
Db	605	TGCACTATGCGCATATGAACAGAGTTGTTCACCGAGATCTCAAGCTGGAAAAATCTTTC	664
QY	682	TAGATGCGCAATGGAACACATCAAGATTTGCTGACTTTGGGCTCTCCAACTGTATACCAAG	741
Db	665	TGATATGCCAATGCGGAATATCAAGATTTGCTGACTTTGGGCTCTCCAACTGTATACCAAG	724
QY	742	GCAAGTTCTCTCAAGCTTTCTGTGGGAGCCCTCTTACGCTCTGCTTGAATGTCAAAG	801
Db	725	GCAAGTTCTCTCAAGCATTTCTGTGGGAGCCCTCTTATGCTCTGCAAGATTTGTCAAATG	784
QY	802	GGAAGCCCTTATGTGGGGCCAGAGGTGAGCAGCTGGTCTCTGGGGGTTTCCCTGTATCACC	861
Db	785	GGAAGCCCTTATCAAGAGCCCAAGAGTGAACAAGTGGTCTCTGGGGTTCCTCTTATCACC	844
QY	862	TGATGCAATGAGCACCATGCGCTTTTGACGGGGAGATCATTAACAATCTGTGTGAAGCAATCA	921
Db	845	TGATGCAATGAGCACCATGCGCTTTTGATGGGAGATGACATTAAGATCTTATGTGAAGCAATCA	904
QY	922	GTAACGGGGCTTACCGTGAAGCCGCCCAAGCGGTCCGATGCTGTGGCTGATCGGTGGC	981
Db	905	GCAACGGGGCTTACCGGGAGGCCACTTAACCTCTGATGCTGTGGCTGATCGGTGGC	964
QY	982	TGTTAATGTGTAAACCCCAACCGCTCGGGGCACATCTGAAGAGATGTAGCCATATGTGTGG	1041
Db	965	TGTTGATGTGTAAACCCCAACCGCTCGGGGCACCTTGAAGAGATGTAGCCATGTGTGG	1024
QY	1042	TCAACTGGGGTTTACACACACCGAGTCTGGGGAAACAGAAAGCCCTGCTGAGGGTGGGCACC	1101
Db	1025	TCAACTGGGGCTTACGCGCACCCGAGTGGGAGAGCAGAGAGGCTTCGATAGGGTGGGCACC	1084
QY	1102	CTAATGTGTGATTTTGGCGCGGGCTTTCATATGGCGGACATGGTTACGTGCTCTTCGCGCCCC	1161
Db	1085	CTGCGAGTGTGATTTTGGCGCGGGCTTTCATATGGCGGACATGGCTCGCGGGTCTCTCCGCCCC	1144
QY	1162	TTCCTGGAAGATGAGAGCCCAAGGTGTGACCTTTTCAACACAGACGCTGCGGGAGGTGGA	1221
Db	1145	TTCCTGGAAGATGAGAGCCCAAGGTGTGACCTTTTCAACACAGATACACTTGGTGGGGAA	1204
QY	1222	GCACTGTACCTGGGCTGGAGCGGCACATTTCTTTAAGAGTCCCGAAGAGAGATGACA	1281
Db	1205	GCACTGTACCTGGGCTGGAGCGGCACATTTGCTTCAAGAGTCCCGAAGAGAGATGACA	1266
QY	1282	TGCGTCAAAAATCTGCAAGGTGACCCCGCTGAGAGATCTCTTTCTGCGCTTGCAAGAGCA	1341
Db	1265	TGCGCCCAAGTCTCTCAACGTGACACAGGCTGTATGACATGCCCATCGCTTGCAAGAGCA	1324
QY	1342	GCGTTAAGCTTCCGAAGGCAATTTCTCAAGAAAAAGTCTCTTACCTCGTCAGGGAGGTAC	1401
Db	1325	ACCTCAAGCTGCCAAGGGCAATTTCTCAAGAAAGAGGTACAGCTCTTCAAGAAAGGGTAC	1384

[illegible]

PR 07-JUL-2000; 2000EP-00114089.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
DR P-PSDB; ADL30886.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
length human cDNAs.
XX
PS Example 1; SEQ ID NO 2918; 1340bp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
Query Match 47.4%; Score 1375.8; DB 12; Length 3395;
Best Local Similarity 81.7%; Pred. No. 4.7e-294;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
QY 82 CTAAGTGTGCGCCGACCACTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 141
DB 77 CTAAGTGTGCGCCGACCACTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 136
QY 142 TCAGAGCGCCGACGAGGCTTCCTCGGCTTCGCGCTTCGAGAGCGCCGCGCGC 201
DB 137 CGGCGCGCTCCGCGCCGACCTCCCTCGGCGGAGAGTCA-----GCCGCGCGC 184
QY 202 TGGCGGAGCGGCTCATGAGTGGCTTAACTCTGAGAGAGAGAGAGAGAGAGAGAGAG 261
DB 185 TGGCGGAG 244
QY 262 ACCATCAACAACAACCTGCGGACCGCTACGAGTTCCTGAGAGAGAGAGAGAGAGAG 321
DB 245 ACCATCAACAACAACCTGCGGACCGCTACGAGTTCCTGAGAGAGAGAGAGAGAGAG 304
QY 322 CTAAGGAG 381
DB 305 CTAAGGAG 364
QY 382 TCAG 441
DB 365 TCAG 424
QY 442 TCAGTGTTCGACCAACCGCCGACCATCTGCGATCTGAGAGAGAGAGAGAGAGAGAG 501
DB 425 TCAGTGTTCGACCAACCGCCGACCATCTGCGATCTGAGAGAGAGAGAGAGAGAGAG 484
QY 502 GCAAGATTTGATTTGATGAGAGATGACGAGCGAGAGAGAGAGAGAGAGAGAGAGAG 561
DB 485 GCAAGATTTGATTTGATGAGAGATGACGAGCGAGAGAGAGAGAGAGAGAGAGAGAG 544
QY 562 AGCGGCGACGCGTGAAG 621
DB 545 AGCGGCGACGCGTGAAG 604
QY 622 TGCACTAGTGCACCAAG 681
DB 605 TGCACTAGTGCACCAAG 664
QY 682 TAGATGCAATGAG 741

DB 665 TGATGCAATGAG 724
QY 742 GCAAGTGTTCGACGAG 801
DB 725 GCAAGTGTTCGACGAG 784
QY 802 GGAAGGCTTATGAG 861
DB 785 GGAAGGCTTATGAG 844
QY 862 TGGTGCATGAG 921
DB 845 TGGTGCATGAG 904
QY 922 GTAAGGCGCTTACCGTGAAG 981
DB 905 GCAAGGCGCTTACCGTGAAG 964
QY 982 TGTATATGAG 1041
DB 965 TGTATATGAG 1024
QY 1042 TCAACTGAG 1101
DB 1025 TCAACTGAG 1084
QY 1102 CTAGTGTGATCTTGGCGGCGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161
DB 1085 CTAGTGTGATCTTGGCGGCGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
QY 1162 TCCTGAG 1221
DB 1145 TCCTGAG 1204
QY 1222 GCACTGTACTGAG 1281
DB 1205 GCACTGTACTGAG 1264
QY 1282 TGGCTCAAAATCTGCAAG 1341
DB 1265 TGGCTCAAAATCTGCAAG 1324
QY 1342 GCTTTAAGCTTCCGAAAG 1401
DB 1325 ACCTCAAGCTTCCGAAAG 1384
QY 1402 AG 1461
DB 1385 AG 1438
QY 1462 TATCCCTGCTCCGAAAG 1521
DB 1439 TATCCCTGCTCCGAAAG 1495
QY 1522 ACTCTCTCCGAG 1581
DB 1496 ACTCTCTCCGAG 1555
QY 1582 GTGGAG 1641
DB 1566 GTGGAG 1615
QY 1642 GCAATTCGAAATCTGCAAG 1701
DB 1616 GCAATTCGAAATCTGCAAG 1675
QY 1702 CTTTGGCTCCCTGAG 1761
DB 1676 CTTTGGCTCCCTGAG 1735
QY 1762 CAGGAG 1821
DB 1736 CAGGAG 1795

QY	1822	TTGAAACGCTCTCCGGAACCCACATGAGGGCTGTGTCTCTGTGACAACTTGAGGGGCG	1881
QY <td>1796</td> <td>CTGAACGGCTCCCAAGGCCCCACATCGCGGGCTGTGTCTGTGTGACAACTCAACGGGGC</td> <td>1855</td>	1796	CTGAACGGCTCCCAAGGCCCCACATCGCGGGCTGTGTCTGTGTGACAACTCAACGGGGC	1855
QY <td>1882</td> <td>TTGACAGACCTCCCTCAGAG-----GTCTGAACGATGTGTGGAGGAAATCCT</td> <td>1929</td>	1882	TTGACAGACCTCCCTCAGAG-----GTCTGAACGATGTGTGGAGGAAATCCT	1929
Db <td>1856</td> <td>TTGAGAGACCCCCCTCAGAGGGCCCTTGAAAGCTGCTGAAGGGCGCTGGCGGAGAGTCCCT</td> <td>1915</td>	1856	TTGAGAGACCCCCCTCAGAGGGCCCTTGAAAGCTGCTGAAGGGCGCTGGCGGAGAGTCCCT	1915
QY <td>1930</td> <td>TGGGGGATAGCTGCTTTTCTCTGACAGACTGCGAAGATGATCTGTGACCTTACAGACAA</td> <td>1989</td>	1930	TGGGGGATAGCTGCTTTTCTCTGACAGACTGCGAAGATGATCTGTGACCTTACAGACAA	1989
Db <td>1916</td> <td>TGGGGGACAGCTGCTTTTCCCTGACAGACTGCGAAGATGATGACAGCACTTACGACAGG</td> <td>1975</td>	1916	TGGGGGACAGCTGCTTTTCCCTGACAGACTGCGAAGATGATGACAGCACTTACGACAGG	1975
QY <td>1990</td> <td>CCCTAGGAATCTGCTCAAGCTACAGCTAGAGAGGAGATGGTCCCTAGTATGGGGTAG</td> <td>2049</td>	1990	CCCTAGGAATCTGCTCAAGCTACAGCTAGAGAGGAGATGGTCCCTAGTATGGGGTAG	2049
Db <td>1976</td> <td>CACAGAGGGTCTGCTCAAAAGCTCACTGATGTGAATGAGCAATGGCCAG--CCGGGTAG</td> <td>2034</td>	1976	CACAGAGGGTCTGCTCAAAAGCTCACTGATGTGAATGAGCAATGGCCAG--CCGGGTAG	2034
QY <td>2050</td> <td>GCTCTGAGAGGGTTTGACAGAGAACCTCTGGGTTCGATTCCTCC</td> <td>2092</td>	2050	GCTCTGAGAGGGTTTGACAGAGAACCTCTGGGTTCGATTCCTCC	2092
Db <td>2035</td> <td>GCTCTCAGATGACCTGGTTGTGACCCCTCAGAGGAGATGCCCTTC</td> <td>2077</td>	2035	GCTCTCAGATGACCTGGTTGTGACCCCTCAGAGGAGATGCCCTTC	2077

RESULT 9
ABD33085
ID ABD33085 standard; cDNA; 3404 BP.

DT 18-NOV-2004 (first entry)

Human cancer-associated (CA) cDNA HR07-004.

KW Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW ss; cancer; cytostatic.

OS Homo sapiens.

PN WO2004058146-A2.

PD 15-JUL-2004.

PF 15-DEC-2003; 2003WO-US040081.

PR 17-DEC-2002; 2002US-00322281.

PA (SAGR-) SAGRES DISCOVERY INC.

PI Morris DW, Malandro MS;

WPI; 2004-499109/47.

XX 5

PT Novel human cancer associated protein encoded within open reading frame of cancer associated gene, useful as targets for diagnosing cancer.

PS Claim 1; SEQ ID NO 25; 182pp; English..

CC The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents human CA cDNA of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published/pct sequences

XX Sequence 3404 BP; 723 A; 1044 C; 895 G; 742 T; 0 U; 0 Other;
SQ

Query Match	47.4%	Score 1375.8;	DB 13;	Length 3404;
Best Local Similarity	81.7%;	Pred. No. 4.7e-294;		
Matches 1652; Conservative	0;	Mismatches 337;	Indels 34;	Gaps 4;

QY	82	TTACTGCTGCGCCGACCCCACTCACTCGCGAGTCCCGGACATGAGAGTGGAGCTTAC	141
Db	87	CTTACTGATTTCCCTGCGCCCTTGTCTCACTCTGCTGCGCATGGAGTGGCTGTTTTCG	146
QY	142	TCAGCGCCCGAGCGCAGGCTCCTCGCGCTCGCCTTGCCCTCGAGAGCGCCCGCCGC	201
Db	147	CGCGCGCTCCGCGCCCACTCCCTCGCGCGAGAGCTA-----GCCCGCGCG	194
QY	202	TGGCGGACGGGCTCATCAAGTCCCTTAACTCTGATGATAAGAGCGCGGTGAAGCGGC	261
Db	195	TGGCGGAAGGGCTGATCAAGTCCGCAAGCCCTTAATGAMAAGCGCGGTGAAGCGGC	254
QY	262	ACCATCACAAAACAACAACCTGCGGACCGGCTACGAGTTCTTGAGACGCTGGGCAAGGCA	321
Db	255	ACCAACAACAACAACAACCTGCGGACCGGCTACGAGTTCTTGAGACCTCGGGCAAAAGCA	314
QY	322	CCTAACGGGAGGTGAAGAGCAAGAGAGCTCGGGCGTCTGCTGGCCATCAAGTCCA	381
Db	315	CCTAACGGGAGGTGAAGAGCGCGGGGAGACTCGGGCGCTCGTGGCCATCAAGTCCA	374
QY	382	TCAGGAAGAACAATCAAAATGATGAGCAGATCTGCTGCAATTCGAGGAGAGATTGGA	441
Db	375	TCGGAGAGCAAAATCAAAATGAGCAGATCTGATGCAATACGAGGAGATTGGA	434
QY	442	TCATGCTTCACTCAACCAACCCCAATCATTTGCCATCATGAGTGTGTAAGATACGA	501
Db	435	TCATGCTCACTCAACCAACCTCAATCATTTGCCATCATGAGTGTGTAAGACAGCA	494
QY	502	GCAAGATTGTGATTGTCAATGAGATATGCGCAACCGAGAGACTGTATATTAATCACTAGT	561
Db	495	GCAAGATTGTATGTCATGAGATATGCGCAACCGGGGGAACCTTATATCACTACATCAGC	554
QY	562	AGCGGCCACGGCTGATGATGAGCGGAGCGCAAGCAATTTCTTCCGACAGATCGTGTGCC	621
Db	555	AGCGGCAGCACTCAGTGAAGCGGCAAGCTAAGCATTTCTTCCGCAAGATCGTCTGCCG	614
QY	622	TGCACTATGCGCACAGAGCGGATCGTTCAACGAGATCTCAAGCTGGAACAATCCTTTC	681
Db	615	TGCACTATGCGCATCAGACAGAGTGTGCCACCGAGATCTCAAGCTGGAACAATCCTCT	674
QY	682	TAGATGCCAATGGAACAATCAAGATGTGTGACTTTGGGCTTCCAACTGTACACAAG	741
Db	675	TGATGCCAATGGAATATCAAGATGTGTGACTTTGGGCTTCCAACTGTACACAAG	734
QY	742	GCAAGTTCCTCAAGCGTTCTGTGGGAGCCCTCTCTACGCTCGCTGAGATAGTCAAG	801
Db	735	GCAAGTTCCTCAAGCATTTGTGGGAGCCCTCTTAAGCTCGCAGAGATTGTCAATG	794
QY	802	GGAAGCCCTAATGTGGGCCCAAGGTGGAACAGTGTCTCTGGGGTTTCTCTGTACATCC	861
Db	795	GGAAGCCCTTACAAGGCCCAAGGTGGAACAGTGTCTCTGGGGTTTCTCTGTACATCC	854
QY	862	TGATGATGAGCAACATGCGCTTTTGAACGGGACGAGATCAATAAACTGGTGAAGCAATCA	921
Db	855	TGATGATGAGCAACATGCGCTTTTGAATGGGACATGACATTAATCTTAATGAAACATATCA	914
QY	922	GTAACGGGGCTTAACCGTGAAGCGCCCAAGCCGTCGATGCTGTGGCTGTATCCGGTGC	981
Db	915	GCAACGGGGCTTAACCGGGAGGCACTTAACCCCTCTGATGCTGTGGCTGTATCCGGTGC	974
QY	982	TGTTAATGATGAACCCCAACCGGTGGGCAACTGAGAGATGTACAGTAAATGATGTGGG	1041
Db	975	TGTTAATGATGAACCCCAACCGGCGGCACTCTGAGAGATGTGCCAGTCACTGTGTGGG	1034
QY	1042	TCAACTGGGGGTAAACCAACCGGAGTCCGGGAAACAGAAAGCCCTGCTAGAGGTGGGCAAC	1101

```
D 1035 TCACTGGGGCTACCCACCCGAGTGGAGACAGAGGCTCCCGCATGAGGCTGGGACCC 1094
Q 1102 CTAGTGTGATCTTTGGCCGGGCTTCATGCGGAGCTGGTTAAGTGGCTCTCGCGCCCC 1161
D 1095 CTGGCAGTGAATCTGGCCGGGCTTCAGTGAATGGCTGGGCTCCCGGCTCCCGCCCC 1154
Q 1162 TCTTGAAGATGAGCCAGAGTGTGACGCTTCTTCAAGCAGACGTCGCGGAGGTGAA 1221
D 1155 TCTTGAAGATGAGGCGCAGAGTGTGACGCTTCTTCAAGCAGACGTCGTCGCGGAG 1214
Q 1222 GCACTGATACCTGGGCTGGAGGGGAGCAATCTTGAAGAACTCCCGAAGAGAAATGA 1281
D 1215 GCACACCCCTGGGCTGGAGGGAGCAATCTTGAAGAACTCCCGAAGAGAAATGA 1274
Q 1282 TGGCTCAAAATCTGCAAGTGAACCCGGCTGAGATACCTTCTTGGCCCTGGCAAGACA 1341
D 1275 TGGCCAGTCTCTCAAGTGAACCCGGCTGAGATACCTTCTTGGCCCTGGCAAGACA 1334
Q 1342 GCCTTAAGCTTCCGAAAGGCAATCTTCAAGAAAGTCTTCAAGTCAAGGAGGTAC 1401
D 1335 ACCTCAAGTCTCCGAAAGGCAATCTTCAAGAAAGTCTTCAAGTCAAGGAGGTAC 1394
Q 1402 AGAGGACCTCTCAGAACTCAGACCGGCTGATATCTCAGGGGAGGCTCTGCTG 1461
D 1395 AGAGGACCTCTCAGAACTCAGACCGGCTGATATCTCAGGGGAGGCTCTGCTG 1448
Q 1462 TATCCCTGCTCCGAAAGGAGCAATCTTGAAGTCTGACAGGCTGAATCTGGTACT 1521
D 1449 ---CCCTGCTCTCCGAAAGGAGCAATCTTGAAGTCTGACAGGCTGAATCTGGTACT 1505
Q 1522 ACTCCTCTCCGAAAGGAGCAATCTTGAAGTCTGACAGGCTGAATCTGGTACT 1581
D 1506 ACTCCTCTCCGAAAGGAGCAATCTTGAAGTCTGACAGGCTGAATCTGGTACT 1565
Q 1582 GTGGGAGCCCTGAGAGCAAGTCTCAGACGCTTCAAGGCTCTCTTCAAGGAGG 1641
D 1566 GTGGGAGTCCGAAAGGAGCAATCTTGAAGTCTGACAGGCTCTCTTCAAGGAGG 1625
Q 1642 GCATTTCAAATCTGAGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1701
D 1626 GCATTTCAAATCTGAGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1685
Q 1702 CTTTGGCTCCCTGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1761
D 1686 CTTTGGCTCCCTGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1745
Q 1762 CAGGGGCTGAGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1821
D 1746 CAGGGGCTGAGAGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1805
Q 1822 CTGAAGCTTCTTCCGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1881
D 1806 CTGAAGCTTCTTCCGCAAGTCTTCCGCAAGCTTGAAGGAGCAATCTTCAAG 1865
Q 1882 TTGAGCAGCTCCCTCAGAGG-----GTTGAGCAGTGTGGAGAGATCT 1929
D 1866 TTGAGCAGCTCCCTCAGAGGCTCCGCAAGTCTTGAAGGAGCAATCTTCAAG 1925
Q 1930 TGGGGAGTACGCTTTTCTTGAAGCTGAGAGGAGTGAAGTGAAGTGAAGTGAAG 1989
D 1926 TGGGGAGTACGCTTTTCTTGAAGCTGAGAGGAGTGAAGTGAAGTGAAGTGAAG 1985
Q 1990 CCCAGAGATCTGTCAGAAAGTCTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2049
D 1986 CACTGAGAGTGTCTCAAGTCTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2044
Q 2050 GCTTGAAGAGGTTTGAAGAGCAAGTGTGGTGGATCTTC 2092
D 2045 GCTTGAAGAGTGTGTGCAAGTGTGGTGGATCTTC 2087
```

RESULT 10
ABX71420

```
ID ABX71420 standard; cDNA, 3443 BP.
XX
AC ABX71420;
XX
DT 14-APR-2003 (first entry)
XX
DE Human cell cycle-associated cDNA from clone DKFZp464j73.
XX
KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
OS Homo sapiens.
XX
PN W0200112659-A2.
XX
PD 22-FEB-2001.
XX
PE 18-AUG-2000; 2000MO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
XX
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEMU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
XX
DR P-PSDB; ABUS3319.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Claim 1, Page 942-943; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence encodes a polypeptide
CC described in the disclosure of the invention
CC
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
Query Match 47.4%; Score 1374.2; DB 5; Length 3443;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Q 82 CTACTGCTGCGCCGACCACTCACTCGCGGTCCCGGACGATGAGTGGCTTAC 141
D 87 CTACTGATTTCCCTGCGCCCTTCTCACTCTGCTCGCATGAGTGGCTTTCG 146
Q 142 TCCAGCGCCGAGCCAGGCTCCCTCGGCTCGGCTTGGCTCGGAGAGCCCGGCGCC 201
D 147 CGCGGCTCTCGGCGCCCACTCTCTCGGCGGAGAGCTA-----GCCGCGCC 194
Q 202 TGGGGAAGGCTCATCAAGTGGCTTAACCTGATGAAGAGAGAGGAGTGAAGCGGC 261
D 195 TGGGGAAGGCTGATCAAGTGGCTTAACCTGATGAAGAGAGAGGAGTGAAGCGGC 254
Q 262 ACCATCAACAACAACCTGAGGAGCGGCTTCAAGTCTTGAAGAGCGTGGGCAAGGCA 321
D 255 ACCATCAACAACAACCTGAGGAGCGGCTTCAAGTCTTGAAGAGCGTGGGCAAGGCA 314
Q 322 CCTACGGGAAGTGAAGAGGAGAGAGAGTGGGCTGCTGAGGAGTGAAGTGAAGTGAAG 381
D 315 CCTACGGGAAGTGAAGAGGAGAGAGAGTGGGCTGCTGAGGAGTGAAGTGAAGTGAAG 374
Q 382 TCAGAAAGACAAATCAAGATGAGAGAGTGTGTCACATACGAGGAGATTTGA 441
D 375 TCCGAAAGACAAATCAAGATGAGAGAGTGTGTCACATACGAGGAGATTTGA 434
```

OY	442	TCATGTCCTCACTCAACACACCCCAATCATATTGGCATCATAGAGGTTTGGAAATGACA	501
Db	435	TCATGTCCTCACTCAACACACCCCAATCATATTGGCATCATAGAGGTTTGGAAATGACA	494
OY	502	GCAAGATTGTGATTTGTCAATGAGATATGCCAGCGAGCGAGCTGTATGATTAATCATCACTG	561
Db	495	GCAAGATTGTGATTTGTCAATGAGATATGCCAGCGCGAGCGAGCTGTATGATTAATCATCACTG	554
OY	562	AGCGGCGACGGGTGAGTGAAGCGGCGACGCAAGGCAATTTCTTCCGACAGATTCGGTCTGCC	621
Db	555	AGCGGCGACAGTTCAGTGAAGCGGCGAGCTTGAAGCAATTTCTTCCGCGAGATTCGTCTGCCG	614
OY	622	TGCACTACTGCCACCAAGAACGGGATCGTTCAACGAGATCTCAAGCTGAGAAAATCATCTCTC	681
Db	615	TGCACTATTGCCATCAGAACAGAGTTGTTCACCGAGATCTCAAGCTGAGAAAATCATCTCTC	674
OY	682	TAGATGCCAATGAGAAACATCAAGATTGCTGACTTTGGGCTCTCCAACTGTACCAAAAG	741
Db	675	TGATGATCCAAATGAGAAATCAAGATTGCTGACTTTGGGCTCTCCAACTGTACCAATCAAG	734
OY	742	GCAAGTTCCTCCAGACGTTCTGTGAGAGCGCCCTCTCAAGCGCTCGCTGAGATGATCAAG	801
Db	735	GCAAGTTCCTCCAGACATTTCTGTGAGAGCGCCCTCTTATGCTCTCGCAAGATTGTCAATG	794
OY	802	GGAAGCCCTATGTGGGCGCCCAAGGTGACAAGTGTCTCTGAGCGCTTCTCTGTACATCC	861
Db	795	GGAAGCCCTTACACAGGCGCCAGAGGTGACAGCTGTCTCCGTGGGTCTTCTCTGTACATCC	854
OY	862	TGGATGATGAGGACCAATGCGCTTTTGAACGGGCAAGGATCATTAACAATGATGAAACAAATCA	921
Db	855	TGGATGATGAGCAACATGCGCTTTTGAATGGCATGACATTAAGATCTTAAGTAAACAGATCA	914
OY	922	GTAACGGGGCTTACCGGTAGCGCCGCAAGCGCTGCATGCTGTGGCTGTATCCGGTGGC	981
Db	915	GCAACGGGGCTTACCGGGAGCCACTTAACCTCTGATGCGCTGTGGCTGTATCCGGTGGC	974
OY	982	TGTTTAATGATGAACCCCAACCCGTGGGGCCACACTGGAAGATGTAACTCATTTGGTGGG	1041
Db	975	TGTTGATGATGAACCCCAACCGCGGGCCACCTGTGAAGATGTGGCCAGTCACTGTGTGG	1034
OY	1042	TCAACTGGGGTTTACACACACCGGAGTGGGGGAAACAGAAAGCCCTGTGTAAGGGTGGGAC	1101
Db	1035	TCAACTGGGGCTACCGCCACCGAGTGGGAGAGACAGAAAGCTTCGATAGGGTGGGAC	1094
OY	1102	CTAAGTGTGACTTTTGGCCGGGCTTCATAGCGGAGCTGTATAGTGCCTTCGCGCCCC	1161
Db	1095	CTGGCAGTGAATCTGCGCCCGCGCTTCATAGGCTGACTGTGCTCGGGGTTCTTCGCGCCCC	1154
OY	1162	TCCGTGGAATATGAGGCCAAGGATGTGACCTTCTTAACACAGACGTGCGGAGGATGAA	1221
Db	1155	TCCGTGGAATATGAGGCCAAGGATGTGACCTTCTTAACACAGATCACTGTGTGGGAA	1214
OY	1222	GCACTGTATCTTGGGTGAGCGGCAACATTTCTTTAAGATGCCGGAAGAGATGACA	1281
Db	1215	GCAACACCCCTGGGCTGTGAGGCGACGATTTGCTTAAGAAATGCCGGAAGAGATGACA	1274
OY	1282	TGGCTCAAAATCTGTCAAGGTGACCCGGCTGAGATCTCTTTCTGCGCTTGCAAGACA	1341
Db	1275	TGGCCCAATCTCTCCACAGTACACCGCTGTATGACATGCCCATTCGCTTGCAAGACA	1334
OY	1342	GCCTTAAGCTTCCGAAGGCAATTTCTCAAGAAAAGTCTCTTACCTCGTCAAGGGAGGATC	1401
Db	1335	ACCTCAAGCTCCGAAGGCAATTTCTCAAGAAAAGTGTCAAGCTTGTGCAAGAGGGATC	1394
OY	1402	AGGAGGACCTTCAGAACTCAGACCGAGTGCCTGTATCTCAAGGCGAGCTGTCCCTGCTG	1461
Db	1395	AGGAGGACCTTCGGAAGCTCAGACCAATCTCTGCGAGGCCAGGGCAGGCTGCC-----	1448
OY	1462	TATCCCTGCTCCCAAGGAAGGCAATCCTTAAGAAATCTCGACAGGTGAATCTGTATCT	1521
Db	1449	---CGCTGCTCCCAAGGAAGGCAATTTCAAGAAAGCCCGACAGGCGCAGTCTGTGCTACT	1505

Oy	1522	ACTCCTCTCCAGAACCCAGACAGTCTGGGGAACTTAAAGACGACAGATGATGTTTGTGA	1581
Oy	1522	ACTCCTCTCCAGAACCCAGACAGTCTGGGGAACTTAAAGACGACAGATGATGTTTGTGA	1581
Db	1506	ACTCCTCTCCAGAACCCAGATGAATCTGGGGAGCTCTTGAACGACGAGCAGTGTTTGTGA	1555
Oy	1582	GTGGGAGACCCGTTGAGACAGAAAGTCTCAACAGCTTCAGAGCTCTCTCTCCACCGCAGG	1641
Db	1566	GTGGGGATATCCCAAGAGACAGAAAGCTCCGCAAGCTTCAGAGGCTCTCTCCATGCAAG	1625
Oy	1642	GCATTTCTCAACATCAATGGCAAGTTCTCCCGCACAAGCTTAAAGAGCATACCCCTAGCA	1701
Db	1626	GCATCTCTCAAACTCAATGGCAAGTTCTCCCGCACAAGCTTAAAGAGCATACCCCTAGCA	1685
Oy	1702	CCTTTGGCTCCCTGAGACCAACTGGCTCTCTCCCTCTGACAGCCGGGCCAGCCGCT	1761
Db	1686	CCTTCGGCTCCCTGAGAGAACTGGCCCCCACTGGCCCCCTGGCCCCGACCCGACCT	1745
Oy	1762	CAGGAGGCTGTGATGAGAGACAGCATCTGTCTTCCTCCAGTCCCTTGAACCAATTTGACTTGC	1821
Db	1746	CAGGAGGCTGTGATGAGAGAGACAGCATCTGTCTTCCTTGAAGCTTGAACCAAGTTGACTTGC	1805
Oy	1822	CTGAACGCTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGACAACTGAGGGGC	1881
Db	1806	CTGAACGCTCTCCAGAACCCCACTGCGGGGCTGTGTCTGTGAGACAACTGAGGGGC	1865
Oy	1882	TTGAGCAGCCCTCCCTCAGAA-----GTCCGAAGCAGATGGTGGCAGGAATCCT	1929
Db	1866	TTTGAAGAGCCCCCTCAGAGGGCCCTGAAAGCTGCTGAAGCGGTGGCGGAGGATCCTT	1925
Oy	1930	TGGGGGATAGCTGTTTCTTCTGTGACAGACTGCCAAGAGTGACTGCACCTACAGACAG	1989
Db	1926	TGGGGGACAGCTGCTTTTCCCTGACAACTGCGCAGAGGATGACAGGACCTACCGACAG	1985
Oy	1990	CCCTAGGAATCTGCTCAAAGCTCAGCTGAGAGAGGAGATGGTCCCTAGTATGGGGTAG	2049
Db	1986	CACGAGAGGTCTGCTCAAACTCACTGAGTGAATGAGCATTTGCCAG-CCGATGCAG	2044
Oy	2050	GCTCTGAGAGGGTTTGCAGAGGAACCTGGGTGGATTCCTCC	2092
Db	2045	GCTCTCAGATGACAGCTGTGCAACCCGAGGGGAGATGCCCTTC	2087

RESULT 11	
ADP76964	
ID	ADP76964 standard; cDNA; 3443 BP.
AC	ADP76964;
XX	
DT	26-FEB-2004 (first entry)
XX	
DB	Novel human secreted and transmembrane protein cDNA SeqID 639.
XX	
KW	human; PRO; membrane bound protein; membrane bound receptor;
KW	cell proliferation; cell migration; cell differentiation;
KW	mitogenic factor; survival factor; cytotoxic factor;
KW	differentiation factor; neuropeptide; hormone; cell receptor;
KW	receptor-ligand interaction; cyostatic; chondrocyte; tumour; ss; gene.
XX	
OS	Homo sapiens.
XX	
PN	WO2003072035-A2.
XX	
PD	04-SEP-2003.
XX	
PF	21-FEB-2003; 2003WO-US005241.
XX	
PR	22-FEB-2002; 2002US-0359461P.
XX	
PA	(GERTH) GENENTECH INC.
XX	
PI	Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
XX	Williams PM, Wood WI, Wu TD;
XX	WPI; 2003-721702/68.

DR P-PSDB; ADF76965.

XX New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
XX diabetes mellitus.

PS Claim 2; SEQ ID NO 639; 918bp; English.

XX
XX This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neurotrophins and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is a cDNA sequence which encodes a human PRO protein of the
XX invention.

SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 47.4%; Score 1374.2; DB 10; Length 3443;

Best Local Similarity 81.6%; Pred. No. 1.1e-293;

Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGTCGCCGACCACTCCACCTCGCGGTCCCGACCATGAGTCGGTGCCTTAC 141
DB 87 CTACTGTCGCCGACCACTCCGCGGTCCCGACCATGAGTCGGTGCCTTAC 146
QY 142 TCCAGCGCCGAGCAGGCTCCCTCGGCTTCGCGCTTCGAGAGCGCCGCGCGC 201
DB 147 CGCGGCGCTCCGCGCCCACTCCCTCGGCGCGAGAGCTA-----GCCGCGCGC 194
QY 202 TGGCGGAGCGGCTCATGATCGCTTAACTCTGATGAGAGAGAGCGGTGAAGGCGC 261
DB 195 TGGCGGAGCGGCTCATGATCGCTTAACTCTGATGAGAGAGAGCGGTGAAGGCGC 254
QY 262 ACCATCAAAACAAACCTCGCGGACCGCTAGAGTTCCTGAGAGCGCTGGGCAAGGCA 321
DB 255 ACCATCAAAACAAACCTCGCGGACCGCTAGAGTTCCTGAGAGCGCTGGGCAAGGCA 314
QY 322 CTTACGGGAAGGTGAAGAGCAAGAGAGCTCGGCGCTCTGTTGGCCATCAAGTCA 381
DB 315 CTTACGGGAAGGTGAAGAGCGCGGAGAGCTCGGCGCTCTGTTGGCCATCAAGTCA 374
QY 382 TCGAGAAAGCAAAATCAAGATGAGAGAGATCTGCTGACATACGAGGAGATTGAGA 441
DB 375 TCGAGAAAGCAAAATCAAGATGAGAGAGATCTGATGACATACGAGGAGATTGAGA 434
QY 442 TCATGCTTCACTCAACACCCCAACATTCATTCGATCATGAGAGTGTGAGAAATGCA 501
DB 435 TCATGCTTCACTCAACACCCCAACATTCATTCGATCATGAGAGTGTGAGAAATGCA 494
QY 502 GCAAGATTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 495 GCAAGATTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 554
QY 562 AGCGGCAACCGCTGATGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 621
DB 555 AGCGGCAACCGCTGATGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 614

QY 622 TGCACTATGCGCACAGAGCGGATGCTTACCGAGATCTCACTGAGAAACTCTCTC 681
DB 615 TGCACTATGCGCACAGAGCGGATGCTTACCGAGATCTCACTGAGAAACTCTCTC 674
QY 682 TAGATGCAATGGAATCAATGATGATGATGATGATGATGATGATGATGATGATG 741
DB 675 TAGATGCAATGGAATCAATGATGATGATGATGATGATGATGATGATGATGATG 734
QY 742 GCAAGTTCCTCAGACGCTTGTGGAGCCCTCTCAAGCTTCGCTGAGATGATGATG 801
DB 735 GCAAGTTCCTCAGACGCTTGTGGAGCCCTCTCAAGCTTCGCTGAGATGATGATG 794
QY 802 GGAAGCCCTATGAGGCGCCAGAGGTGACAGCTGCTTGGGCGCTTCTCTGATCACC 861
DB 795 GGAAGCCCTATGAGGCGCCAGAGGTGACAGCTGCTTGGGCGCTTCTCTGATCACC 854
QY 862 TGGTGCATGACCAATGCGCTTGTGACGGGAGAGATCAATTAACCTGATGAGAAATCA 921
DB 855 TGGTGCATGACCAATGCGCTTGTGATGAGGAGATGACATGATCTGATGAGAAATCA 914
QY 922 GTAAAGGAGCTTACCGTGAAGCGCGCCCAAGCGCTGATGCTGATGCTGATGCTG 981
DB 915 GCAAGGAGGCTTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 974
QY 982 TGTATATGATGAACCCCAAGCGGCGCACTGAGAGATGATGATGATGATGATGATG 1041
DB 975 TGTATATGATGAACCCCAAGCGGCGCAAGCGGCGCAAGCGGCGCAAGCGGCGCA 1034
QY 1042 TCAACTGAGGCTTACCAACCGGAGTGGGAGCAAGAGAGCGCTGAGAGGCTGAGG 1101
DB 1035 TCAACTGAGGCTTACCAACCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1094
QY 1102 CTAGTGTGATCTTGGCGGCGGCGCTCAATGAGGAGCTGATGATGATGATGATGATG 1161
DB 1095 CTAGTGTGATCTTGGCGGCGGCGCTCAATGAGGAGCTGATGATGATGATGATGATG 1154
QY 1162 TCTGAGAGATGAGAGCGGAGTGTGAGCTTCTTCAAGCAGACAGTGGCGGAGGTGAA 1221
DB 1155 TCTGAGAGATGAGAGCGGAGTGTGAGCTTCTTCAAGCAGACAGTGGCGGAGGTGAA 1214
QY 1222 GCACTGATCTGAGGCTGAGAGCGGAGCAATTTCTTGAAGAGTCCCGAAGAGATGACA 1281
DB 1215 GCACTGATCTGAGGCTGAGAGCGGAGCAATTTCTTGAAGAGTCCCGAAGAGATGACA 1274
QY 1282 TGGCTCAAAATCTGCAAGTGTGACCGGCTGAGATCTCTTCTGCGCTGCGCAAGACA 1341
DB 1275 TGGCTCAAAATCTGCAAGTGTGACCGGCTGAGATCTCTTCTGCGCTGCGCAAGACA 1334
QY 1342 GCTTAAAGCTTCCGAAAGGAGATTTCTCAAGAAAGAGTCTCTTCAAGGAGGATGAC 1401
DB 1335 GCTTAAAGCTTCCGAAAGGAGATTTCTCAAGAAAGAGTCTCTTCAAGGAGGATGAC 1394
QY 1402 AGAGAGACCTCAGAGAACTGACCGGCTGCTGATCTCAAGGAGGCTGCTGCTGCTG 1461
DB 1395 AGAGAGACCTCAGAGAACTGACCGGCTGCTGATCTCAAGGAGGCTGCTGCTGCTG 1448
QY 1462 TATCCCTGCTCCCAAGAGAGGAGTCTCTTAAAGAGTCTGCAAGCGGAGATCTGTTACT 1521
DB 1449 ---CGTGTCTCCCAAGAGAGGAGTCTCTTAAAGAGTCTGCAAGCGGAGATCTGTTACT 1505
QY 1522 ACTCTCTCAGAGAGCGGAGGAGTCTGAGGAACTTGAAGCGGAGGATGATGATGATG 1581
DB 1506 ACTCTCTCAGAGAGCGGAGGAGTCTGAGGAACTTGAAGCGGAGGATGATGATGATG 1565
QY 1582 GTGGGAGCCCGTGGAGCAAGAGTCTTCAAGGCTTCAAGGCTTCTCTTCAAGGAGG 1641
DB 1566 GTGGGAGTCCCAAGAGCAAGAGCTTCCGAGAGCTTCAAGGCTTCTCTTCAAGGAGG 1625
QY 1642 GCATTCTCAAACTCAATGAGGAGGATCTCCGCAAGGCTTGAAGAGGAGGAGGAGGAG 1701
DB 1626 GCATTCTCAAACTCAATGAGGAGGATCTCCGCAAGGCTTGAAGAGGAGGAGGAGGAG 1685
QY 1702 CTTTGGCTCCCTGAGCAAACTGGGCTCTTCCATCTGAGAGCGGCGGCGGCGCT 1761

Db 1686 CTTTGGCTCCCTGGATGAATCGCCCACTCCGCCCCCGGCGCCAGCCGCT 1745
 Qy 1762 CAGGGGCTGTGATGAGAGACGATCTGTCTCCGAGTCTTTGACCAATGACCTTGC 1821
 Db 1746 CAGGGGCTGTGATGAGAGACGATCTGTCTCCGAGTCTTTGACCAATGACCTTGC 1805
 Qy 1822 CTGAAGCTTTCCCGAAGACCCCACTGAGGGGCTGTGTCTGTGAGCAACTGAGGGGC 1881
 Db 1806 CTGAAGCTTTCCCGAAGACCCCACTGAGGGGCTGTGTCTGTGAGCAACTGAGGGGC 1865
 Qy 1882 TTGAGCAGCTCCCTCAGAGG-----GTCTGAAGCGATGTGTGAGAGATCTT 1929
 Db 1866 TTGAGAGAGCCCTCCAGAGGCCCCCTGAACTGCTGAGGCGCTGGCGGAGATCTT 1925
 Qy 1930 TGGGGGATAGCTGCTTTTCTCTGACAGCTGCCAAGAGTGACTGACGCTTACAGACAG 1989
 Db 1926 TGGGGGACAGCTGCTTTTCTCTGACAGCTGCCAAGAGTGACTGACGCTTACAGACAG 1985
 Qy 1990 CCCTAGGAATCTGCTCAAGCTCAGCTGAGAGAGAGATGTGCTCCCTAGATGAGGATAG 2049
 Db 1986 CACTGAGGGTCTGCTCAAGCTCAGCTGAGAGAGATGTGCTCCCTAG-CCGGTCTAG 2044
 Qy 2050 GCTCTGAGAGGGTTTGGCAGAGAACCTTGGGTCCGATTCTCC 2092
 Db 2045 GCTCTGAGATGCAAGCTGTGCACTCCGAGGGAGATGCTTCC 2087

RESULT 12
 ADF81952
 ID ADF81952 standard; DNA; 3443 BP.
 AC ADF81952;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Leukemia-related DNA sequence #2508.
 XX
 KW Cytostatic; Gene therapy; leukemia; ss.
 OS Unidentified.
 XX
 PN WO2003039443-A2.
 PD 15-MAY-2003.
 XX
 PF 04-NOV-2002; 2002WO-EP012303.
 XX
 PR 05-NOV-2001; 2001EP-00126244.
 PR 30-APR-2002; 2002EP-00009758.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
 PA (HAFE/) HAFERLACH T.
 PA (SCHO/) SCHOCH C.
 PA (KERN/) KERN W.
 XX
 PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schultze S, Dugas M,
 PI Eils R, Biers B, Mergenthaler S;
 DR WPI; 2003-505037/47.
 XX
 PT Determining the subtype of leukemia cells and whether a patient sample
 PT contains leukemia cells or other cells, useful for treating leukemia,
 PT comprises determining the expression profile of a group of markers in a
 PT patient sample.
 XX
 PS Disclosure; SEQ ID NO 2508; 2938bp; English.
 CC The present invention relates to a method (M1) for determining the
 CC subtype of leukemia cells and whether a patient sample contains
 CC leukemia cells. The method comprises determining the expression profile
 CC of a group of markers in a patient sample. The method is useful for

CC determining the presence of leukemia cells, its types or subtypes, and
 CC for the preparation of a medicament for treating leukemia.
 XX
 SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
 Query Match 47.4%; Score 1374.2; DB 10; Length 3443;
 Best Local Similarity 81.6%; Pred. No. 1.1e-293;
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
 Qy 82 CTACTGTGCGCCGACCTTCACCTTCGCGGCTCCGACCATGAGTGGCTTAC 141
 Db 87 CTACTGATTCCTCCGCGCCCTTGCTCAGCTCCGCTCGCATGAGATCTGCTTTG 146
 Qy 142 TCCAGCCGCCGAGCGACGCTTCCTCGGCTTCGCGCTTGGCTTGGAGAGCGCCGCGC 201
 Db 147 CGGCGCCCTCCGCGCCCACTCCCTCGGCGCCAGAGTAA-----GCCCGCGCG 194
 Qy 202 TGGCGAGCGGGCTCATGATCGCTTAACTCTGATGAGAGAGAGCGGGTGAAGCGG 261
 Db 195 TGGCGAGAGGCTGATCAAGTCCGCAAGCTCCCTAATGAGAGAGCGGGTGAAGCGG 254
 Qy 262 ACCATCAAAACAAACCTGCGGACCGCTTACGATTCCTGAGAGCGCTGGAGAGCA 321
 Db 255 ACCACCAAAACAAACCTGCGGACCGCTTACGATTCCTGAGAGCGCTGGAGAGCA 314
 Qy 322 CTTACGGGAAAGTGAAGAGGCAACGAGAGCTTGGGGCTCTGTGTGCTCATCAATCCA 381
 Db 315 CTTACGGGAAAGTGAAGAGGCGCGGAGAGCTTGGGGCTCTGTGTGCTCATCAATCCA 374
 Qy 382 TCAAGAAAGCAAAATCAAGATGAGAGGATCTGCTGCACTTACGAGGAGGATTTGAGA 441
 Db 375 TCCGAAAGCAAAATCAAGATGAGAGGATCTGCTGCACTTACGAGGAGGATTTGAGA 434
 Qy 442 TCAATCTTCTCACTCAACCAACCCCACTATTCGATTCATCAAGTGTGAGAAATGCA 501
 Db 435 TCAATGATCATCTCAACCAACCCCACTATTCGATTCATCAAGTGTGAGAAATGCA 494
 Qy 502 GCAAGATTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
 Db 495 GCAAGATTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
 Qy 562 AGCGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
 Db 555 AGCGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
 Qy 622 TGCATCTACTGCAACGAGAGCGGATGCTTCAACGAGATCTCAAGCTGAGAAACATCTTC 681
 Db 615 TGCATCTACTGCAACGAGAGCGGATGCTTCAACGAGATCTCAAGCTGAGAAACATCTTC 674
 Qy 682 TAGATGCCAATGGAACATCAAGATGCTGATCTTGGCTCTCCAACTGTACCAAG 741
 Db 675 TAGATGCCAATGGAACATCAAGATGCTGATCTTGGCTCTCCAACTGTACCAAG 734
 Qy 742 GCAAGTCCCTCCAGAGCTTGTGGAGAGCCCTCTCAAGCTGCGCTGAGATGATGACAG 801
 Db 735 GCAAGTCCCTCCAGAGCTTGTGGAGAGCCCTCTCAAGCTGCGCTGAGATGATGACAG 794
 Qy 802 GGAAGCCCTATGTGGGCCAGAGATGAGACAGTGTCTCTGGGCGTTCTCTGATACATCC 861
 Db 795 GGAAGCCCTATGTGGGCCAGAGATGAGACAGTGTCTCTGGGCGTTCTCTGATACATCC 854
 Qy 862 TGGTGATGGACACATGCTCTTTGACGGGAGGATCATAAACATGTGTGAAGCAATCA 921
 Db 855 TGGTGATGGACACATGCTCTTTGATGGGATGATCAATGATCTGTGTAACATCA 914
 Qy 922 GTAAAGGGGCTTACGATGAGAGCGGCGCAAGCGCTTCCATGCTGTGCTGTATCCGGTGC 981
 Db 915 GTAAAGGGGCTTACGATGAGAGCGGCGCAAGCGCTTCCATGCTGTGCTGTATCCGGTGC 974
 Qy 982 TGTAAATGTGAACCCCAACCGTGGGCGCACTGAGAGATGATGACAGTATGTGGG 1041
 Db 975 TGTAAATGTGAACCCCAACCGTGGGCGCACTGAGAGATGATGACAGTATGTGGG 1034

QY 1042 TCAACTGGGGTAAACCAACCGGAGTCGGGGAACAGAAACCTCGTGAGGGGTGGACCC 1101
 DB 1035 TCAACTGGGGTAAACCAACCGGAGTCGGGGAACAGAAACCTCGTGAGGGGTGGACCC 1094
 QY 1102 CTAGTGTGATCTTGGGCGGGGCTCAATGGGCGAGCTGGTTAGTGTGCTCGGCGCC 1161
 DB 1095 CTGGGAGTGAATCTGGCGGCGCTCAATGGGCGAGCTGGTTAGTGTGCTCGGCGCC 1154
 QY 1162 TCCTGAGAAATGAGCAAGGTCGACGCTTCTCAAGCAGACAGTGCAGGAGGTGGA 1221
 DB 1155 TCCTGAGAAATGAGGAGGTCGACGCTTCTCAAGCAGACAGTGCAGGAGGTGGA 1214
 QY 1222 GCACTGATCTGGGCTGAGCGGCAACATCTTCAAGAGTCCCGAAGAGAAATGCA 1281
 DB 1215 GCACCAACCCCTGGGCTGAGCGGCAACATCTTCAAGAGTCCCGAAGAGAAATGCA 1274
 QY 1282 TGGGTCAAAATCTGCAAGGTACCCGGGTGAGATCCTCTTTCGCGCTTGGCAAGACA 1341
 DB 1275 TGGGCGAGTCTCTCAAGTACACGCTGATGACATGCCCATCGCCCTGGCAAGACA 1334
 QY 1342 GCCTTAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTTCACTCGTCAAGGAGGTAC 1401
 DB 1335 ACCTCAAGCTCCCAAGGCAATCTCAAGAAAGTGTGAGCTCTGCAAGAGGGTAC 1394
 QY 1402 AGAGAGACCTCAGAGAACTCAGACCGGTGCTGATCTCAAGGAGGCTGCTGCTG 1461
 DB 1395 AGAGAGACCTCAGAGAACTCAGACCGGTGCTGATCTCAAGGAGGCTGCTGCTGCTG 1448
 QY 1462 TATCCCTGCTCCCAAGAGGCAATCTTCAAGAAAGTCTGCAAGGAGGCTGCTGCTG 1521
 DB 1449 ---CGTGTCTCCCAAGAGGCAATCTTCAAGAAAGTCTGCAAGGAGGCTGCTGCTG 1505
 QY 1522 ACTCCTCTCCAGAGCCGAGAGTCTGGGGAATCTTGAAGCCAGTGAATGTTTGA 1581
 DB 1506 ACTCCTCTCCAGAGCCGAGAGTCTGGGGAATCTTGAAGCCAGTGAATGTTTGA 1565
 QY 1582 GTGGGGAACCCGAGAGGAGAGTCTCAGAGGCTTCAAGGCTCTCTCCAGCCGACAG 1641
 DB 1566 GTGGGGAATCCCAAGAGGAGAGTCTCAGAGGCTTCAAGGCTCTCTCCAGCCGACAG 1625
 QY 1642 GCATTCTCAAACTCAATGGCAAGTCTTCCGCAAGGCTTGAAGGCACTAACCCCTGCA 1701
 DB 1626 GCATCTCAAACTCAATGGCAAGTCTTCCGCAAGGCTTGAAGGCACTAACCCCTGCA 1685
 QY 1702 CCTTGGCTCCCTGGAACCACTGAGCTCTCCCAATCTGAGCCGCGCCAGCCCT 1761
 DB 1686 CCTTGGCTCCCTGGAATGACTGCGCCCACTCGCCCTGCGCCGCGCCAGCCCT 1745
 QY 1762 CAGGCGCTGTGAGTGAAGCAGATCTGTCTCCGAGTCTTGAACCAATTTGACTTGC 1821
 DB 1746 CAGGCGCTGTGAGTGAAGCAGATCTGTCTCCGAGTCTTGAACCAATTTGACTTGC 1805
 QY 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGCGCTGTGTCTGTGACAACTGAGGCGGC 1881
 DB 1806 CTGAAGCGCTCTCCAGAGCCCACTGAGGCGCTGTGTCTGTGACAACTGAGGCGGC 1865
 QY 1882 TTGAGCAGCCTCCCTCAGAAAG-----GTGCAACGATGTGTGAGAGAAATCCT 1929
 DB 1866 TTGAGAGAGCCCTCCAGAGGCGCTGAGAGCTGCTGAGGCGCTGAGGAGAGATCCTT 1925
 QY 1930 TGGGAGATAGTGTCTTCTTCTGACAGACTGCCAAGAGGTACTGACAGCTTACAGACAG 1989
 DB 1926 TGGGAGATAGTGTCTTCTTCTGACAGACTGCCAAGAGGTACTGACAGCTTACAGACAG 1965
 QY 1990 CCCTGAAGATCTGCTCAAAAGCTCAGCTGAGAGAGGAGAGTGTGCTTATGTGGGTAG 2049
 DB 1986 CACTGAGAGGTGTGCTCAAAAGCTCAGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2044
 QY 2050 GCTCTGAGAGGCTTTCAGAGAGAACCTGAGGTGAGATTCCTCC 2092
 DB 2045 GCTCTGAGATCAGCTGTGAGTGTGAGTGTGAGGAGAGATTCCTTC 2087

RESULT 13
 ID ADO20171
 AD ADO20171 standard; cDNA; 3443 BP.
 XX
 AC ADO20171;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human PRO polynucleotide #540.
 XX
 KW Human; PRO; gene; ss; immune related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KW renal disease; demyelinating disease; central nervous system;
 KW peripheral nervous system; demyelinating polynuropathy;
 KW Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polynuropathy.
 XX
 OS Homo sapiens.
 XX
 PN WO2004043361-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 06-NOV-2003; 2003WO-US035268.
 XX
 PR 08-NOV-2002; 2002US-0425235P.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX
 DR WPI; 2004-420067/39.
 XX
 DR P-PSDB; ADO20172.
 XX
 PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO8388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 XX
 PS Claim 1; SEQ ID NO 1136; 1731bp; English.
 XX
 CC The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polynuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polynuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.
 XX
 SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
 XX
 Query Match 47.4%; Score 1374.2; DB 12; Length 3443;
 Best Local Similarity 81.6%; Pred. No. 1.1e-293;
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
 QY 82 CTACTGTGCGGACCACTCAGCTGCGGTGCTCCCGCACTGAGTGTGCTTAC 141
 DB 87 CTACTGATTCCTCCGCGGCTTCTCACTCTGCTGCGCATGAGTGTGCTTTCG 146
 QY 142 TCCAGGCGCGGAGCAGGCTCTCCGCGCTCCGCGCTGCGAGAGCGCCGCGCGC 201
 DB 147 CGGCGGCTCCGCGCCCACTCTCCGCGCGAGGCTA-----GCCGCGCGCGC 194
 QY 202 TGGCGAGCGGCTCATCAAGTGTGCTTAACCTTGATGAGAGACAGGCGGTGAAGCGGC 261

Db 195 TGGGGAAGGCTATCAAGTCGCGCAAGCCCTTATATAAGAGCGGCTGAAGCGC 254
Qy 262 ACCATCACAAAACACAACTGCGGCAACCGCTACGAGTTCTGGAAGCGTGGGCAAGGCA 321
Db 255 ACCACCAAGACAAACCTGCGGCAACCGCTACGAGTTCTGGAAGCGTGGGCAAGGCA 314
Qy 322 CCTACGGGAAGGTAAAGAGGCAAGAGAGCTCGGGGCGTGTGGCCATCAAGTCA 381
Db 315 CCTACGGGAAGGTAAAGAGGCGCGGAGAGCTCGGGGCGTGTGGCCATCAAGTCA 374
Qy 382 TCAGAAAGACAAATCAAAAGATGAGAGATCTGCTGCATACATACGAGGAGATTGGA 441
Db 375 TCCGGAAGGACAAATCAAAAGATGAGAGATCTGATCAATACGAGGAGATTGGA 434
Qy 442 TCATGCTTCACTCAACCAACCCCAATCATTTGCCATCATGAAGTGTGGAATGCA 501
Db 435 TCATGCTATCACTCAACCAACCTCAATCATTTGCCATCATGAAGTGTGGAAGCA 494
Qy 502 GCAAGATTGTGATGTCAATGAGATATGCGACCGAGCGGAGTGTATATTAATCAATG 561
Db 495 GCAAGATTGTGATGTCAATGAGATATGCGACCGAGCGGAGCTTTATGATCAATCA 554
Qy 562 AGCGGCAAGGCTGATGAGAGCGGAGCGGAGCAATTTCTCGGACAGATCGTGTGCG 621
Db 555 AGCGGCAAGGCTGATGAGAGCGGAGCGGAGCAATTTCTCGGAGAGTGTGTGCG 614
Qy 622 TGCACTACGCGACCAAGACGGAGTCTTCAACGAGATCTCAAGCTGGAATAATCTTC 681
Db 615 TGCACTATTTGCTATCAAGACAGAGTTTGCACCGAGATCTCAAGCTGGAATAAT 674
Qy 682 TAGATGCAATGGAACATCAAGATGTGATCTTTGGCTCTCCACCTGATCAACAA 741
Db 675 TAGATGCAATGGAACATCAAGATGTGATCTTTGGCTCTCCACCTGATCAACAA 734
Qy 742 GCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTTACGCTCGCTGAGATGATGCAAG 801
Db 735 GCAAGTCTCTCAAGCATTTCTGTGGAGCCCTCTTATGCTCTCGCAGAGATTTG 794
Qy 802 GGAAGCCCTATGTGGGCGGAGAGTGAACAGCTGTCTTGGGCGTTCTCTGTACAT 861
Db 795 GGAAGCCCTACACAGGCGGAGAGTGAACAGCTGTCTTGGGCGTTCTCTGTACAT 854
Qy 862 TGTGTGATGGAACCATGCTCTTTGAAGGAGATCATTAACATGTGTGAAGCAATCA 921
Db 855 TGTGTGATGGAACCATGCTCTTTGAAGGAGATCATTAACATGTGTGAAGCAATCA 914
Qy 922 GTAACGGGCTTACGATGAGCGGCGCAAGCGGTGCGATGCTGTGGCTGTATCGG 981
Db 915 GCAAGCGGCTTACGAGGAGGCACTTAACCTCTGTATGCTGTGTGCTGTATCGG 974
Qy 982 TGTATATGTGAACCCCAACCGGTGGGCAACATGAGAGATGAGCCAGTATGTGG 1041
Db 975 TGTATATGTGAACCCCAACCGGCGGCAACCTGAGAGATGTGGCCAGTCACTGT 1034
Qy 1042 TCAACTGGGGTTACCAACCGGAGTGGGGAACAGAAACCTTGGTGAAGGGTGG 1101
Db 1035 TCAACTGGGGTTACCAACCGGAGTGGGGAACAGAAACCTTGGTGAAGGGTGG 1094
Qy 1102 CTAGTGTGATCTTTGGCGGGGCTCAATGGGAGCTGGTTAAGTGTGCTCTCGGCG 1161
Db 1095 CTAGTGTGATCTTTGGCGGGGCTCAATGGGAGCTGGTTAAGTGTGCTCTCGGCG 1154
Qy 1162 TCCTGAGAAATGAGCAAGGTGTGAGCTTTCTTCAAGCAACGTCGCGGAGGTG 1221
Db 1155 TCCTGAGAAATGAGGAGCAAGGTGTGAGCTTTCTTCAAGCAACGTCGTCGAG 1214
Qy 1222 GCACTGATCTGGGCTGAGGCGCAATTTCTTTAAGAGTCCCGAAAGAGATGACA 1281
Db 1215 GCACTGATCTGGGCTGAGGCGCAATTTCTTTAAGAGTCCCGAAAGAGATGACA 1274
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTCTTCTCGGCTGGCAAGCA 1341

Db 1275 TGGCCAGTCTCTCCACAGTACACAGGCTATGACATCGCCATCGCCCTGGCAAGACA 1334
Qy 1342 GCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCTTCACTGTGAGGAGTAC 1401
Db 1335 ACCTCAAGCTCCCAAGAGGCAATTTCTCAAGAAAAGTGTGAGCTTGTGCAAGAGG 1394
Qy 1402 AGAGGACCTCCAGGAACTGAGACCGGTGTATATCTCAGGGCACTGTCTCTGTG 1461
Db 1395 AGAGGACCTCCAGGAACTGAGACCGGTGTATATCTCAGGGCACTGTCTCTGTG 1448
Qy 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTTAAAGATCTGCAACAGCTGAATCTG 1521
Db 1449 ---GCTGCTCCCAAGAGGCAATTTCTCAAGAAACCCCAAGCGGAGCTGTGACT 1505
Qy 1522 ACTCTCTCCAGAGCCCAAGAGTCTGGGAACTTTAAGACCGCATGATGTGTTGGA 1581
Db 1506 ACTCTCTCCAGAGCCCAAGGAAATCTGGGAGCTTTGAGACGAGGGAAGTGTGGA 1565
Qy 1582 GTGGGACCCCGTGGAGCAAGATCTGCAACAGCTTCAAGGCTCTCTCAACCGCAAG 1641
Db 1566 GTGGGAGTCCCAAGAGCAAGAGCTCCGCAAGCTTCAAGGCTGTCTCTCATGCAAG 1625
Qy 1642 GCATTCTCAAACTCAATGGCAAGTTCTCCGCAACAGCTTAAAGGCACTAACCCCTGCA 1701
Db 1626 GCATCTCAAACTCAATGGCAAGTTCTCCGCAACAGCTTGAAGCTGTGCGGCCCA 1685
Qy 1702 CTTTGGCTCTCTGAGCAACTGGCTCTTCCATCTGCAAGCCGAGCCCAAGCGCCCT 1761
Db 1686 CTTTGGCTCTCTGAGTGAAGTGGCTCTGCAAGCTGTGCTGTGCTGTGCAACCTCA 1745
Qy 1762 CAGGGGCTGTGAGAGAGACAGATCTGTCTCTCCAGTCTTTGACCAATTTGACTTGC 1821
Db 1746 CAGGGGCTGTGAGAGAGACAGATCTGTCTCTCTGAGTCTTTGACCAAGCTTGC 1805
Qy 1822 CTGAACGCTTCTCCGAAACCCCACTGAGGGGCTGTGTCTGTGCAACACTGAGGGG 1881
Db 1806 CTGAACGCTTCTCCGAGAGCCCACTGCGGGCTGTGTCTGTGCAACACTCAAGGG 1865
Qy 1882 TTGAGCAGCTCTCCCTCAAGG-----GTCTGAAGCAATGTGTGGCAGGAATCT 1929
Db 1866 TTGAGGAGCCCTCCCTCAAGAGGCTGTGAAGGCTGTGAAGGCTGTGGGAGATCT 1925
Qy 1930 TGGGGATATGCTGTTTCTGTGACATCTGCAAGAGTGTGCTGACGCTTACAGACAG 1989
Db 1926 TGGGGATATGCTGTTTCTGTGACATCTGCAAGAGTGTGCTGACGCTTACAGACAG 1985
Qy 1990 CCTGAGGATCTGCTCAAGGCTGAGTGAAGAGGAGATGTGCTCTAATATGGGGTAC 2049
Db 1986 CACTGAGGATCTGCTCAAGGCTGAGTGAAGAGTGTGCTCTAATATGGGGTAC 2044
Qy 2050 GCTCTGAGAGGTTTTCAGAGGAACCTGTGGATCGATTCTCC 2092
Db 2045 GCTCTGAGATGCAAGCTGTGTGACCCGAGGGGAGATGCTTTC 2087

RESULT 14
ADJ96554
ID ADJ96554 standard; DNA; 3463 BP.
XX
AC ADJ96554;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human calcium/calmodulin-dependent protein kinase Mnak2 DNA Segid 11.
XX
KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW calcium/calmodulin-dependent protein kinase; Mnak2.
XX
OS Homo sapiens.

OS 68. Location/Qualifiers
FH Key replace(1670,c)
FT variation /*tag= a
FT /*standard_name= "Single nucleotide polymorphism"
FT /*tag= b
FT /*tag= b
FT /*standard_name= "Single nucleotide polymorphism"
PN MO2004006838-A2.
PD 22-JAN-2004.
PE 15-JUL-2003; 2003WO-US021730.
PF 15-JUL-2002; 2002US-0395632P.
PR 15-JUL-2002; 2002US-0395632P.
PX (SUGEN-) SUGEN INC.
PY Whyte D, Manning G, Caenepeel S;
PI WPI: 2004-122753/12.
DR P-PSDB; ADJ96620.
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX Example 1; SEQ ID NO 11; 366bp; English.
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTKs and STKs),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytoprotective, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
XX
SQ Sequence 3463 BP; 732 A; 1057 C; 919 G; 755 T; 0 U; 0 Other;
Query Match 47.4%; Score 1374.2; DB 12; Length 3463;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
QY 82 CTACTGCTGCCGACCACTCACTCGCGGTCCCGACCAATGAGTCGGCTTAC 141
DB 148 CTATTGATTCCTCCGCGCCCTTGTCTCACTCTGCTCCATGAGTCGGCTTTCG 207
QY 142 TCGACGCGCCGAGCGAGGCTCCTCGGCTCGCCCTCGGAGCGCGCGCGCGC 201
DB 208 CGCGGCGCTCCGCGCCCACTCTCGGCGCGAGACTA-----GCCGCGCGC 255
QY 202 TGGCGAGCGGCTCATCAAGTCGCTTAACTCTGATGAGAGAGCGGCTGAAGCGGC 261
DB 256 TGGCGAGAGGGCTGATCAAGTCGCGCAAGCCCTTAATGAAGAGCGGCTGAAGCGGC 315
QY 262 ACCATCAAAACAACCTCGCGGACCGGCTAGAGTTCTTGAGACGCTGGGCAAGGCA 321
DB 316 ACCACCAAAACAACCTCGCGGACCGGCTAGAGTTCTTGAGACGCTGGGCAAGGCA 375
QY 322 CCTACGAGAGGTGAAGAGGACGAGAGAGCTCGGCGCTCTGCTGAGCAATCAAGTCA 381
DB 376 CCTACGAGAGGTGAAGAGGCGGAGAGAGCTCGGCGCGCTGCTGAGCAATCAAGTCA 435
QY 382 TCAGGAAAGCAAAATCAAGATGAGCAAGATCTGCTCAATACGAGGAGGATTTGA 441

DB 436 TCCGAAAGCAAAATCAAGATGAGCAAGATCTGATGCAATACGAGGAGATTGGA 495
QY 442 TCATGCTTTCATCAACACCCCAATCATTTGCTCATCAATGAAGTTTGAATAGCA 501
DB 496 TCATGCTTTCATCAACACCCCAATCATTTGCTCATCAATGAAGTTTGAATAGCA 555
QY 502 GCAAGATTGATTTGTCATGAGATGATGACCGGAGGAGATCTGATGATTAATCACT 561
DB 556 GCAAGATTGATTTGTCATGAGATGATGACCGGAGGAGATCTGATGATTAATCACT 615
QY 562 AGCGGCAACCGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
DB 616 AGCGGCAACCGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675
QY 622 TGCACTACTGCGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
DB 676 TGCACTACTGCGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 735
QY 682 TAGATGCCAATGAGAAATCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
DB 736 TAGATGCCAATGAGAAATCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 795
QY 742 GCAAGTCTCTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
DB 796 GCAAGTCTCTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855
QY 802 GGAAGCCCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
DB 856 GGAAGCCCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
QY 862 TGGTGCATGAGCAACATGCTTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
DB 916 TGGTGCATGAGCAACATGCTTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
QY 922 GTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
DB 976 GTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
QY 982 TGTAAATGAGTAAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
DB 1036 TGTAAATGAGTAAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095
QY 1042 TCAACTGAGGAGTAAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
DB 1096 TCAACTGAGGAGTAAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1155
QY 1102 CTAGTGTGACTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
DB 1156 CTAGTGTGACTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1215
QY 1162 TCTGAGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
DB 1216 TCTGAGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1275
QY 1222 GCACTGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1281
DB 1276 GCACTGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335
QY 1282 TGGCTCAAAATTCGAAGAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1341
DB 1336 TGGCTCAAAATTCGAAGAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1395
QY 1342 GCTTAAAGCTTTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1401
DB 1396 GCTTAAAGCTTTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1455
QY 1402 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1461
DB 1456 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
QY 1462 TATCCCTGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1521
DB 1510 ---CCTGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1566

QY 1522 ACTCTCTCCAGAGCCGAGAGTCTGGGGAACTCTTAGAGCCAGTGAATGTTTNGA 1581
DB 1567 ACTCTCTCCAGAGCCGAGAGTCTGGGGAACTCTTAGAGCCAGTGAATGTTTNGA 1626
QY 1582 GTGGGGAGCCCGTGGAGAGAGTCTCCAGAGCTTCAAGGGCTCTCTCCAGCGAAG 1641
DB 1627 GTGGGGATCCCAAGAGAGAGAGCTCCGAGAGTTCAAGGGCTCTCTCATGCGAAG 1686
QY 1642 GCATTCTCAAACTCAATGGCAAGTTCTCCGCAAGCCTTAGAAGCACTACCTTAGCA 1701
DB 1687 GCATCTCTCAAACTCAATGGCAAGTTCTCCGAGAGCTTGAAGCTCCGCGCCCAACA 1746
QY 1702 CCTTGGCTCCCTGAGCAACTGGAGCTCTCCAGTCCGAGCCCGCGCCAGCCGCT 1761
DB 1747 CCTTGGCTCCCTGAGCAACTGGAGCTCTCCAGTCCGAGCCCGCGCCAGCCGCT 1806
QY 1762 CAGGGGCTGTGAGTGAAGAGAGCATCTGTCTCCAGTCCCTTGAACCAATTGACTTGC 1821
DB 1807 CAGGGGCTGTGAGTGAAGAGAGCATCTGTCTCCAGTCCCTTGAACCAAGTTCG 1866
QY 1822 CTGAAGCTCTTCCGAAACCCCACTGAAGGGCTGTGTCTGTGAGCAACTGAGGGGC 1881
DB 1867 CTGAAGCTCTTCCGAAACCCCACTGAAGGGCTGTGTGTGTGTGAGCAACTGAGGGGC 1926
QY 1882 TTGAGCAGCCTCCCTCAGAGG-----GTGGAAGGAGTGTGGAGGAATCCT 1929
DB 1927 TTGAGAGAGCCCTCCAGAGGGCCCTGAGAGCTCTGAGGGCTGGGGGAGGATCCTT 1986
QY 1930 TGGGGGATGAGTGTCTTCTCTGACAGCTGCGCAAGAGTGAAGTCAAGCTTACAGAGAG 1989
DB 1987 TGGGGGAGAGCTGCTTTTCCCTGACAGCTGCCAGAGGTGAAGAGCACTTACCGAGAG 2046
QY 1990 CCTTGAAGATCTGTCTCAAGCTCAGCTGAGAGAGAGAGTGTGCTTATGAGGGTAG 2049
DB 2047 CACTGAGAGGCTGTCTCAAGCTCAGCTGAGTGAAGTGAAGTGTGCTTACCGAG-CCGGGTACG 2105
QY 2050 GCTCTGAGAGGTTTGAAGAGAACCTGGGTCCGATTTCTCC 2092
DB 2106 GCTCTGAGTGAAGTGTGACCCCGAGGAGAGATGCTTC 2148

RESULT 15

ID ADM43851 standard; cDNA; 2501 BP.
XX ADM43851;
AC ADM43851;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #215.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.
PR 20-AUG-2002; 2002US-00225251.
XX
PA (TANG/) TANG Y T.
PA (XUE/) XUE A.
PA (DRMA/) DRMANAC R T.
XX
PI Tang YT, Xue A, Drmanac RT;
XX

DR WPI: 2004-238579/22.
XX
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
XX useful for diagnosing and/or treating conditions associated with aberrant
XX activity of the arginine-rich polypeptides, such as cancer and
XX inflammation.
XX
XX Disclosure; SEQ ID NO 215; 51pp; English.
XX
XX The invention relates to an isolated polynucleotide. The methods and
XX compositions of the present invention are useful for the diagnosis and/or
XX treatment of diseases or conditions associated with aberrant expression
XX or activity of the arginine-rich protein-like polypeptides, such as
XX cancer and inflammation. They can also be used in forensics, gene
XX mapping, identification of mutations responsible for genetic disorders,
XX and in assessing biodiversity. The present sequence represents a novel
XX human arginine-rich protein cDNA.
SQ Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;

Query Match 47.3%; Score 1372.2; DB 12; Length 2501;
Best Local Similarity 81.6%; Pred. No. 2,7e-293;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTGCCAGCCACTCCACCTGCGCGGTCCCGCAGCAGTGAAGTCCGCTTAC 141
DB 65 CTATTGATTCCTCCCTGCGCCCTTGTCTACCTCTGCTGCGCATGAGTCTGCTGTTTCG 124
QY 142 TCCAGCCGCCGAGCCAGAGCTCCCTCGGCTCCGCTCGGCTCGAGAGCCCGGCGCG 201
DB 125 GGGGGGCTCCGGGCCACTCCCTCGGCTCCGAGAGCTA-----GCCCGGCGCG 172
QY 202 TGGGGGAGGGCTTATCAATGCTGCTTAACTTGTATGAAGAGAGCGGTGAAGCGGC 261
DB 173 TGGGGGAGGGCTGATCAATGCTGCTTAACTTGTATGAAGAGAGCGGTGAAGCGGC 232
QY 262 ACCATCAACAACAACCTGCGGCAACGCTTGAAGTCTTGAAGAGCGGTGAAGAGGCA 321
DB 233 ACCATCAACAACAACCTGCGGCAACGCTTGAAGTCTTGAAGAGCGGTGAAGAGGCA 292
QY 322 CTTACGGGAGAGTGAAGAGGCAAGAGAGCTCGGGGCGCTGTGGTGGCCATCAAGTCA 381
DB 293 CTTACGGGAGAGTGAAGAGGCGGGAGAGCTCGGGGCGCTGTGGTGGCCATCAAGTCA 352
QY 382 TCAAGAAAGCAAAATCAAGATGAGCAGGATCTGTGCAATACGAGGAGAGATTGAGA 441
DB 353 TCCGAAAGAGCAAAATCAAGATGAGCAGGATCTGTGCAATACGAGGAGAGATTGAGA 412
QY 442 TCAATGCTTCACTCAACACCCCACTCACTTGCATCCATGAGTGTGAGAAATGCA 501
DB 413 TCAATGCTTCACTCAACACCCCACTCACTTGCATCCATGAGTGTGAGAAATGCA 472
QY 502 GCAAGATTGTGATTTGTATGATGATGCGAGCGAGCGAGTGTATGATTAATCATCAGTG 561
DB 473 GCAAGATTGTGATTTGTATGATGATGCGAGCGAGCGAGTGTATGATTAATCATCAGTG 532
QY 562 AGCGGCCAGCGCTGAGTGAAGCGGAGCGCAGGATTTCTTCCGACAGATGTCGTCGCC 621
DB 533 AGCGGCCAGCGCTGAGTGAAGCGGAGCGCAGGATTTCTTCCGACAGATGTCGTCGCC 592
QY 622 TGCATCTATGCGCACAGAGAGGATGTTTCAAGAGATCTCAAGCTGAGAAACATCTTTC 681
DB 593 TGCATCTATGCGCACAGAGAGGATGTTTCAAGAGATCTCAAGCTGAGAAACATCTTTC 652
QY 682 TAGATGCCAATGGAATCATCAAGATTGCTGATTTGGCTCTCAACCTGTACCAAGAG 741
DB 653 TAGATGCCAATGGAATCATCAAGATTGCTGATTTGGCTCTCAACCTGTACCAAGAG 712
QY 742 GCAAGTTCCTCAAGAGTTCGTGAGAGCCCTCTTACAGCTCGCTGAGATGATCAACG 801
DB 713 GCAAGTTCCTCAAGAGTTCGTGAGAGCCCTCTTACAGCTCGCTGAGATGATCAACG 772
QY 802 GGAAGCCCTATGTGGGCCAGAGAGTGAAGCAGTGTCTGTGGCGCTTCTCTGTATCATCC 861

Db 773 GGAAGCCCTACACAGGCGCCAGAGGTGAGCAGCTGCTCCCTGAGGTTCTCTCTACATCC 832
Qy 862 TGGTGCATGGACCAATCCCTTTTGAACGGGAGATCATTAACAACATGATGAGCAATCA 921
Db 833 TGGTGCATGGACCAATCCCTTTTGAACGGGAGATCATTAACAACATGATGAGCAATCA 892
Qy 922 GTAACGGGAGCTTACCGTGAACCGGCGCAAGCGTCCGATGCTGTGAGCTGTATCCGATGC 981
Db 893 GCAACGGGAGCTTACCGGAGCGCACTTAACCGCTGTATGCTGTGAGCTGTATCCGATGC 952
Qy 982 TGTATATGTGAAACCCCAACCGTGTGAGCTGTATGAGATGTATGAGATGATGAGG 1041
Db 953 TGTATATGTGAAACCCCAACCGGCGCAACCGTGTGAGATGTATGAGATGATGAGG 1012
Qy 1042 TCAACTGGGGTTACACACCGGAGTCCGGGAGACAGAAACCGTGTGAGGAGTGGAGCC 1101
Db 1013 TCAACTGGGGTTACACACCGGAGTCCGGGAGAGAGAGAGGCTCCGCAATGAGGATGGAGCC 1072
Qy 1102 CTAGTGTGACTTTGGCCGGGCTTCATGGCGGACTGTGATTAAGTGTGCTCTCCGCGCC 1161
Db 1073 CTGGCAGTGACTTGCCCGGCGCTTCATGGCTGATGCTGCTCCGCGCTTCCTCCGCGCC 1132
Qy 1162 TCTTGAAGATGAGACCGCAAGGTGTGAGCTTCTTCAAGCAGCAGTGTCCGGAGGTGAA 1221
Db 1133 TCTTGAAGATGAGACCGCAAGGTGTGAGCTTCTTCAAGCAGCAGTGTGTGGGAAA 1192
Qy 1222 GCACTGTACCTGGGCTGAGACCGGAGCAATCTCTTAAGAAAGTCCGAAAGGAGATGACA 1281
Db 1193 GCAACACCGCTGGCTGAGAGCGCAATCTCTTAAGAAAGTCCGCAAGGAGATGACA 1252
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTCTTCTGCGCTGCGCAAGACA 1341
Db 1253 TGGCCAGTGTCTTCCACAGTGAACCGGCTGATGACATGCGCCATGCGCTGCGCAAGACA 1312
Qy 1342 GCTTAAAGCTTCGAAAGGCAATCTCAAGAAAAGTCTCTTCACTGCTCAAGGAGGTAC 1401
Db 1313 ACTCAAGCTGCGCAAGGCAATCTCAAGAAAAGTGTGAGCTCTGCAAGAGGGTAC 1372
Qy 1402 AGGAGACCTCTCAGGAATCTGAGACCGGCTGTATCTCAGGAGCACTGTCCCTGCTG 1461
Db 1373 AGGAGACCTCTCAGGAATCTGAGACCGGCAATCTCTGAGACCGGCAAGGCTGCTC 1426
Qy 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTCTTAAGAAAGTCTGACAGCTGAAATCTGTTACT 1521
Db 1427 ---CCCTGCTCCCAAGGAAAGGCAATCTCAAGAAAGCCCAAGCGGAGTCTGCTACT 1483
Qy 1522 ACTCTCTCAGAGCCCGAGGAGTCTGAGGAACTTTAAGCGCAGTGAATGTGTTTGA 1581
Db 1484 ACTCTCTCAGAGCCCGAGGAGTCTGAGGAGTCTTGAAGCGCAGGCGAGCTGTTTGA 1543
Qy 1582 GTGGGAGCCCGTGTGAGACAGATCTCAACAGGCTTCAAGGCTCTCTCTCAACGCAAG 1641
Db 1544 GTGGGAGTCCCAAGAGACAGAGCTCGCAAGCTTCAAGGCTCTCTCTCATGCAAG 1603
Qy 1642 GCATTTCAAACTCAATGAGCAAGTTCTCCGCAAGGCTTGAAGGCACTAACCCCTAGCA 1701
Db 1604 GCATTTCAAACTCAATGAGCAAGTTCTCCGCAAGGCTTGAAGGCTTGAAGGCTTGAAG 1663
Qy 1702 CTTTGGCTCCCTGAGCAAACTGGCTCTCTCCATCTGACAGCCCGGCGCAAGCCCT 1761
Db 1664 CTTTGGCTCCCTGAGCAAACTGGCTCCCACTGGCCCGCTGGCCCGGCGCAAGCCCT 1723
Qy 1762 CAGGGGCTGTGAGTGAAGACAGATCTGTCTCTCCAGTCTTGAACCAATGAGCTTGC 1821
Db 1724 CAGGGGCTGTGAGTGAAGACAGATCTGTCTCTCCAGTCTTGAACCAATGAGCTTGC 1783
Qy 1822 CTGAACGTCTTCCCGAAACCCCACTGAGGAGCTGTGTCTGTGAACAACCTGAGGAGGC 1881
Db 1784 CTGAACGTCTTCCCGAAACCCCACTGAGGAGCTGTGTCTGTGAACAACCTGAGGAGGC 1843
Qy 1882 TTGAGCAGCTCTCTCAAGAG-----GTCTGAAGCAATGTTGAGGAGAAATCTT 1929

Db 1844 TTGAGAGCCCCCTCAGAGAGGCGCTGGAAGCTGCTGAGAGCGCTGGCGCAGATCTT 1903
Qy 1930 TGGGGATGAGCTGCTTCTCTGACAGACTGCCAAGAGTGAAGTGAAGCTTACAGACAAG 1989
Db 1904 TGGGGATGAGCTGCTTCTCTGACAGACTGCCAAGAGTGAAGTGAAGCTTACAGACAAG 1963
Qy 1990 CCTTGAATCTGCTCAAACTGAGCTGAGAGAGGAGATGAGTCCCTAGTATGAGGTAG 2049
Db 1964 CACTGAGGAGTCTGCTCAAACTGAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2022
Qy 2050 GCTTGAAGAGGTTTGAAGAGAAACCTGGAGTGAATCTT 2090
Db 2023 GCTTCAAGTGAAGCTGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 2063

Search completed: May 15, 2005, 03:21:42
Job time : 1520.61 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 06:02:36 ; Search time 12445.6 Seconds
(without alignments)
11298.520 Million cell updates/sec

Title: US-09-980-464-4
Perfect score: 2902
Sequence: 1 cactagctgagctcccaagaat.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rtc:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	6	AR279568 Sequence
2	2902	100.0	2902	6	AR562226 Sequence
3	2820.4	97.2	2938	10	BC046833 Mus muscu
4	2763.8	92.5	2917	10	BC033302 Mus muscu
5	2102.6	72.5	2900	10	BC081899 Rattus no
6	2066.8	71.2	2929	6	AX380960 Sequence
7	2022.2	69.7	2027	6	AX381046 Sequence
8	1771.4	61.0	2026	6	AX380986 Sequence
9	1593.2	54.9	291762	2	AC125887 Rattus no
10	1379	47.5	3353	6	AX407116 Sequence
11	1375.8	47.4	3395	6	CQ782778 Sequence
12	1375.8	47.4	3395	6	BD127309 Primer fo
13	1375.8	47.4	3395	9	AK074830 Homo sapi
14	1375.8	47.4	3402	9	BC017306 Homo sapi
15	1374.2	47.4	3443	6	AX086949 Sequence
16	1374.2	47.4	3443	6	AX780351 Sequence
17	1370.6	46.9	3360	6	AX642961 Sequence
18	1360.2	46.9	1884	6	AX407118 Sequence
19	1360	46.9	1887	6	CQ714327 Sequence

20	1349.2	46.5	2291	6	AX056394 Sequence
21	1278.6	44.1	252978	2	AC126983 Rattus no
22	1230	42.4	1833	6	AX803413 Sequence
23	1071.4	36.9	1186	6	AX380972 Sequence
24	943.4	32.5	1186	6	AX381029 Sequence
25	849.2	29.3	2616	6	AX430369 Sequence
26	849.2	29.3	2619	6	AX430468 Sequence
27	732	25.2	179414	9	AC098933 Homo sapi
28	728.8	25.1	6389	9	HSMB06181 Homo sapi
29	675.4	23.3	741	6	AX380991 Sequence
30	651.8	22.5	1274	5	BX950299 Gallus ga
31	648	22.3	707	6	AX381050 Sequence
32	648	22.3	712	6	AX380990 Sequence
33	644.2	22.2	661	6	AX381052 Sequence
34	643	22.2	734	6	AX381051 Sequence
35	583.2	20.1	588	6	AX381056 Sequence
36	557.4	19.2	660	6	AX380992 Sequence
37	534	18.4	578	6	AX381054 Sequence
38	515	17.7	548	6	AX381053 Sequence
39	513	17.7	521	6	AX381053 Sequence
40	511.6	17.6	579	6	AX380994 Sequence
41	507.6	17.5	6825	6	CQ715220 Sequence
42	507.6	17.5	6828	6	CQ76374 Sequence
43	507.6	17.5	6828	6	AX213156 Sequence
44	507.6	17.5	6828	9	AB011109 Homo sapi
45	506.4	17.5	1986	6	AX213158 Sequence

ALIGNMENTS

RESULT 1
AR279568
LOCUS AR279568
DEFINITION Sequence 4 from patent US 6514719.
ACCESSION AR279568
VERSION AR279568.1 GI:29714427
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2902)
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.
TITLE Methods for identifying compounds that alter kinase activity
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;
FEATURES
source
1..2902
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2902; DB 6; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACTAGTGATCCAAAGATTCCGACGAGCGGTCTGGGTGGCTGTGACTGTAG 60
DB 1 CACTAGTGATCCAAAGATTCCGACGAGCGGTCTGGGTGGCTGTGACTGTAG 60
QY 61 CCCGCGGCTGACGCGCGGCTGCTACTGCTGCGCCGACCACTCCACTGCGGCTCCGCGCA 120
DB 61 CCCGCGGCTGACGCGCGGCTGCTACTGCTGCGCCGACCACTCCACTGCGGCTCCGCGCA 120
QY 121 CCATGAGTGGTGGTGGCTTACTCCAGGCGCCGACGAGCTCCTCGGCTCCGCGCTGG 180
DB 121 CCATGAGTGGTGGTGGCTTACTCCAGGCGCCGACGAGCTCCTCGGCTCCGCGCTGG 180
QY 181 CCTGAGAGAGCCCGGCGGCTGCGAGCGGCTCATCAAGTCCCTAACTGTATGA 240
DB 181 CCTGAGAGAGCCCGGCGGCTGCGAGCGGCTCATCAAGTCCCTAACTGTATGA 240
QY 241 AGAAGCAGCGCGGTGAACCGGACCATCAACAACAACTGCGGACCGCTACGAGTTCC 300
DB 241 AGAAGCAGCGCGGTGAACCGGACCATCAACAACAACTGCGGACCGCTACGAGTTCC 300

Db	241	AGAGCAGGCGGTGAAGGGGACCATTCACAAACACACTGGGGACCGCTACAGATTCC	300
QY	301	TGGAGACGCTGGGCAAGGGCACCTTACGGGAGGTGAAGAGGCACGAGAGACTGGGGC	360
Db	301	TGGAGACCTCTGGGCAAGGGCACCTTACGGGAGGTGAAGAGGCACGAGAGACTGGGGC	360
QY	361	GTCGTGGTGCATCAAGTCCATCAGAGAAAGACAAATTAAGATAGACAGATCTGCTGC	420
Db	361	GTCGTGGTGCATCAAGTCCATCAGAGAAAGACAAATTAAGATAGACAGATCTGCTGC	420
QY	421	ACATACGAGGAGATTTGAGATCATGTCTTCACTCAACCAACCCGACATCTTGGCATTC	480
Db	421	ACATACGAGGAGATTTGAGATCATGTCTTCACTCAACCAACCCGACATCTTGGCATTC	480
QY	481	ATGAAGTGTTTGAGAAATGACGCAAGATTGTGATCGATGATGCGCACGCCGAGCG	540
Db	481	ATGAAGTGTTTGAGAAATGACGCAAGATTGTGATCGATGATGCGCACGCCGAGCG	540
QY	541	ATCTGATGATTTACATTCAGTAGCGGGCCACCGGCTGAGTGAAGGGGACGCGAGCATTTCT	600
Db	541	ATCTGATGATTTACATTCAGTAGCGGGCCACCGGCTGAGTGAAGGGGACGCGAGCATTTCT	600
QY	601	TCCGACAGATGCTGTCTGCCCTGACCTACTGCGCACAGAACGGAGATCGTTACCGAGATC	660
Db	601	TCCGACAGATGCTGTCTGCCCTGACCTACTGCGCACAGAACGGAGATCGTTACCGAGATC	660
QY	661	TCGAAGCTGAAAACATCCTTTCTAGATGCGCAATGAGAAACATCAAGATTGCTGATTGGCC	720
Db	661	TCGAAGCTGAAAACATCCTTTCTAGATGCGCAATGAGAAACATCAAGATTGCTGATTGGCC	720
QY	721	TCTCCAACTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGAGCCCTCTCTACG	780
Db	721	TCTCCAACTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGAGCCCTCTCTACG	780
QY	781	CCTGCGCTGAAATAGTCAACGGGAAAGCCCTTATGTGGGCGCAAGGCTGACAGTGTCTC	840
Db	781	CCTGCGCTGAAATAGTCAACGGGAAAGCCCTTATGTGGGCGCAAGGCTGACAGTGTCTC	840
QY	841	TGGGCGTTCTCTGTACATCTGTGTGATGGCACCATGCCCTTTGACGGGACGATCATTA	900
Db	841	TGGGCGTTCTCTGTACATCTGTGTGATGGCACCATGCCCTTTGACGGGACGATCATTA	900
QY	901	AAACACATGTGTGAAGAAATCAATGTAACGGGGGCTTACCCTGAGCGGCCCAAGCGTCCGATG	960
Db	901	AAACACATGTGTGAAGAAATCAATGTAACGGGGGCTTACCCTGAGCGGCCCAAGCGTCCGATG	960
QY	961	CCTGTGGCCTGATCCGGTGGGCTGTTAATGTGTGAACCCGACCCGTGGGGCCACATGGAGG	1020
Db	961	CCTGTGGCCTGATCCGGTGGGCTGTTAATGTGTGAACCCGACCCGTGGGGCCACATGGAGG	1020
QY	1021	ATGTAGCCAGTCATTGTGTGGTCAACTGGGGGTTACACCAACGGAGTCGGGAAACAGAAAG	1080
Db	1021	ATGTAGCCAGTCATTGTGTGGTCAACTGGGGGTTACACCAACGGAGTCGGGAAACAGAAAG	1080
QY	1081	CCCTGCGGTGAGGGGTGGGACCCCTGATGTGTGACTTTGGCGCGGACCTCAATGCGGACTGCT	1140
Db	1081	CCCTGCGGTGAGGGGTGGGACCCCTGATGTGTGACTTTGGCGCGGACCTCAATGCGGACTGCT	1140
QY	1141	TACGTGCTCTCTCGGGCCCTCTCTCTGAGAAATGAGGCCAAGGTGTGACGCTTCTTCAAGC	1200
Db	1141	TACGTGCTCTCTCGGGCCCTCTCTCTGAGAAATGAGGCCAAGGTGTGACGCTTCTTCAAGC	1200
QY	1201	AGCAGCTGTCCGGGAGGTGGAGCACTGTACCTGGGCTGAGACGGGCAACATTCTTTAAGA	1260
Db	1201	AGCAGCTGTCCGGGAGGTGGAGCACTGTACCTGGGCTGAGACGGGCAACATTCTTTAAGA	1260
QY	1261	AGTCCCGAAAGGAATGACATGCGTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCT	1320
Db	1261	AGTCCCGAAAGGAATGACATGCGTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCT	1320
QY	1321	CTTCTGCGCCCTGGGAAAGGACGCTTTAAGCTTTCCGAAAGGCAATTCTCAAGAAAAGTCTT	1380
Db	1321	CTTCTGCGCCCTGGGAAAGGACGCTTTAAGCTTTCCGAAAGGCAATTCTCAAGAAAAGTCTT	1380

OY	1381	CTACTCGTCAGGGAGGGATGACAGAGGAACTCTCAGGAACTCAGACCGGGTCCGTGATCTC	1440
Db	1381	CTACTCTGTCAAGGGAGGGATGACAGAGGAACTCTCAGGAACTCAGACCGGGTCCGTGATCTC	1440
OY	1441	CAGGGCAGCTGTCTCCCTGTATCTCTGTCTCCAGGAAAGGCATCTTTAAGAGTCTC	1500
Db	1441	CAGGGCAGCTGTCTCCCTGTATCTCTGTCTCCAGGAAAGGCATCTTTAAGAGTCTC	1500
OY	1501	GACACCGGAACTGTGGTTACTCTCTCTCAGAGCCGACGAGTCTGGGAACTCTTAG	1560
Db	1501	GACACCGGAACTGTGGTTACTCTCTCTCAGAGCCGACGAGTCTGGGAACTCTTAG	1560
OY	1561	ACGCCAGTGAATGTGTTTGGATGGGGAAACCCGTGGAGGAGAGTCTCCACAGGCTTCAG	1620
Db	1561	ACGCCAGTGAATGTGTTTGGATGGGGAAACCCGTGGAGGAGAGTCTCCACAGGCTTCAG	1620
OY	1621	GCGTCTCTCCACCGCAAGGCATCTTCAAACTCAATGGCAATTCTCCGACAGCCCT	1680
Db	1621	GCGTCTCTCCACCGCAAGGCATCTTCAAACTCAATGGCAATTCTCCGACAGCCCT	1680
OY	1681	TAGAAAGGCACTACCCCTAGCACCTTTGGTCCCTGGACCACTGGCTCTCTCCATCTTG	1740
Db	1681	TAGAAAGGCACTACCCCTAGCACCTTTGGTCCCTGGACCACTGGCTCTCTCCATCTTG	1740
OY	1741	CAGCCCGGCCAGCCGCCCTCAGGGGCTGTAGTGAAGACAGCATCTGTCCGAGT	1800
Db	1741	CAGCCCGGCCAGCCGCCCTCAGGGGCTGTAGTGAAGACAGCATCTGTCCGAGT	1800
OY	1801	CCTTTGACCAATGGACTGTGCTGTAACGTCTTCCGAAACCCCACTGAGGGGCTGTGT	1860
Db	1801	CCTTTGACCAATGGACTGTGCTGTAACGTCTTCCGAAACCCCACTGAGGGGCTGTGT	1860
OY	1861	CTGTGGAACAACCTGAGGGGGCTTGAAGCAGCTCCTCAGAAAGTCTGAAGCCATGGTGGC	1920
Db	1861	CTGTGGAACAACCTGAGGGGGCTTGAAGCAGCTCCTCAGAAAGTCTGAAGCGATGGTGGC	1920
OY	1921	AGGAATCCTTGGGGGATAGCTGCTTTTCTCTAGACGCTGCCAAGAGTGACTGCAGCCT	1980
Db	1921	AGGAATCCTTGGGGGATAGCTGCTTTTCTCTAGACGCTGCCAAGAGTGACTGCAGCCT	1980
OY	1981	ACAGCAAGCCCTAGAGAACTGTCTCAAAAGCTCAGCTGAGGAAAGGAGATGTCCTAGT	2040
Db	1981	ACAGCAAGCCCTAGAGAACTGTCTCAAAAGCTCAGCTGAGGAAAGGAGATGTCCTAGT	2040
OY	2041	ATGGGGTAGGCTCTGAAGAGGGTTTCCAGAGAAACCTGGGTGGATTCTCCAGTGAATA	2100
Db	2041	ATGGGGTAGGCTCTGAAGAGGGTTTCCAGAGAAACCTGGGTGGATTCTCCAGTGAATA	2100
OY	2101	GAGTACATCAAGGGCTCTACGCTGTCAGCCTBACTGAACCTGAAAGATAGAGAAATCGC	2160
Db	2101	GAGTACATCAAGGGCTCTACGCTGTCAGCCTBACTGAACCTGAAAGATAGAGAAATCGC	2160
OY	2161	ATTGATGTGAAAGGAATGGGAACCTTGTCTGCTCCGAGTGTATAGTGGGGTGCCTGAA	2220
Db	2161	ATTGATGTGAAAGGAATGGGAACCTTGTCTGCTCCGAGTGTATAGTGGGGTGCCTGAA	2220
OY	2221	GGTGCTTACCTCTTTGTGCCATAGTGTCAACCATTGACATTTCCACCTGTCTCTGG	2280
Db	2221	GGTGCTTACCTCTTTGTGCCATAGTGTCAACCATTGACATTTCCACCTGTCTCTGG	2280
OY	2281	CTGCACCTTCACATTAAGTTTCTGTTTCCATCAACACAGGGTTTGAACCTCGACTTCT	2340
Db	2281	CTGCACCTTCACATTAAGTTTCTGTTTCCATCAACACAGGGTTTGAACCTCGACTTCT	2340
OY	2341	GGAGAGTATGTATGATGACTGCCATTTATAGAGAGAAACAGCCTGTGTTTCCATCT	2400
Db	2341	GGAGAGTATGTATGATGACTGCCATTTATAGAGAGAAACAGCCTGTGTTTCCATCT	2400
OY	2401	CTGCTGTGTGTGATCTCAAAAGCTCTGGAAAGCTGGACCGCTGTTTGACTTCAATCTCA	2460
Db	2401	CTGCTGTGTGTGATCTCAAAAGCTCTGGAAAGCTGGACCGCTGTTTGACTTCAATCTCA	2460

Db	1321	CTTCCTCGCCCTGCGAAGGACGACCTTAAGCTTCGAAAGGCAATTCTCAAGAAAAAGTCTC	1380
OY	1381	CTACCTCGTCAAGGGAGGTACAGAGAGACCTCAGAGAACTGACACGGTGCTGTATATC	1440
Db	1381	CTACTCGTCAAGGGAGGTACAGAGAGACCTCAGAGAACTGACACGGTGCTGTATATC	1440
OY	1441	CAGGGCAACCTGTCCCTCTGTATATCCCTGTCTCCCAAGSAAAGGATCTCTTAAGAAGTTC	1500
OY	1501	GACAGCGTGAATCTGTGTACTACTCCTCTCCAGAGACCCGAGGAGTCTGGGGAATCTTGA	1560
Db	1501	GACAGCGTGAATCTGTGTACTACTCCTCTCCAGAGACCCGAGGAGTCTGGGGAATCTTGA	1560
OY	1561	ACGCCAGTATGTGTGTGTGAATGGGGACCCCGTGGAGCAGAAGTCTTCACAGGCTTCAG	1620
Db	1561	ACGCCAGTATGTGTGTGTGAATGGGGACCCCGTGGAGCAGAAGTCTTCACAGGCTTCAG	1620
OY	1621	GGCTCTCTCTCCACCGGCAAGGGGCAATCTTCAAACTCAATGGCAGTCTTCCGCAAGCTT	1680
Db	1621	GGCTCTCTCTCCACCGGCAAGGGGCAATCTTCAAACTCAATGGCAGTCTTCCGCAAGCTT	1680
OY	1681	TAGAAAGGCACTACCCCTAGCAACCTTTGGCTCCGTGGACCAATCGGCGCTCCATCCGTG	1740
Db	1681	TAGAAAGGCACTACCCCTAGCAACCTTTGGCTCCGTGGACCAATCGGCGCTCCATCCGTG	1740
OY	1741	CAGCCCGGCGCAGCGCCCTCAGGGGCTGTGAATGAGACAGCATCTGTCTTCCAGT	1800
Db	1741	CAGCCCGGCGCAGCGCCCTCAGGGGCTGTGAATGAGACAGCATCTGTCTTCCAGT	1800
OY	1801	CCTTTGACCAATTGGAATCTTGCTGTAAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTGT	1860
Db	1801	CCTTTGACCAATTGGAATCTTGCTGTAAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTGT	1860
OY	1861	CTGTGSAACAACCTGAGGGGGGCTTGAGGACCCCTCCAGAAAGTCTGAAGGAGTGTGTGC	1920
Db	1861	CTGTGSAACAACCTGAGGGGGGCTTGAGGACCCCTCCAGAAAGTCTGAAGGAGTGTGTGC	1920
OY	1921	AGGAATCCTTGGGGATAGCTGTCTTTCTCTGACAGACTGACCAAGGCTGACTCAGCT	1980
Db	1921	AGGAATCCTTGGGGATAGCTGTCTTTCTCTGACAGACTGACCAAGGCTGACTCAGCT	1980
OY	1981	ACAGACAAGCCCTTAGAATCTGTCTCAAAAGCTCAGCTGAGAAAGGAGATGTGTCCCTAGT	2040
Db	1981	ACAGACAAGCCCTTAGAATCTGTCTCAAAAGCTCAGCTGAGAAAGGAGATGTGTCCCTAGT	2040
OY	2041	ATGGGGTAGGCTCTGAGAGGGGTTTGACAGAGAACCTCGGGTCCGATTCCTCAATGAATA	2100
Db	2041	ATGGGGTAGGCTCTGAGAGGGGTTTGACAGAGAACCTCGGGTCCGATTCCTCAATGAATA	2100
OY	2101	GAGTACATCAAGGGCTCTACGCTCTGACGCTGACTGAACCTGAAACCTGAGAGAAATGCG	2160
Db	2101	GAGTACATCAAGGGCTCTACGCTCTGACGCTGACTGAACCTGAAACCTGAGAGAAATGCG	2160
OY	2161	ATTGATGTGAAAGGAATGGGAACCTTGTGCTGCCCGAGTGTATATGTGGGTGCGCTGAA	2220
Db	2161	ATTGATGTGAAAGGAATGGGAACCTTGTGCTGCCCGAGTGTATATGTGGGTGCGCTGAA	2220
OY	2221	GGTGCCTAACCTCTTTGTGSCATAGTGTACCCATGACATTTTCCACCCCTGTCTCTGG	2280
Db	2221	GGTGCCTAACCTCTTTGTGSCATAGTGTACCCATGACATTTTCCACCCCTGTCTCTGG	2280
OY	2281	CTGCACCTTCAATTAAGTTTCTGTGTTCCATCAACCAAGGGGTTAGAACCTGTGACTTCT	2340
Db	2281	CTGCACCTTCAATTAAGTTTCTGTGTTCCATCAACCAAGGGGTTAGAACCTGTGACTTCT	2340
OY	2341	GGAGAGTATGTGTAGTACTGCACTTAATTTAGAGAGGAAACAGCTCTGTGTTTCAATCT	2400
Db	2341	GGAGAGTATGTGTAGTACTGCACTTAATTTAGAGAGGAAACAGCTCTGTGTTTCAATCT	2400
OY	2401	CTGCTGTGTGCAATCTCAAAAGCCTGGGAAAGCTGGAACCGCTGTGTAATTCAATCTCAA	2460
Db	2401	CTGCTGTGTGCAATCTCAAAAGCCTGGGAAAGCTGGAACCGCTGTGTAATTCAATCTCAA	2460

Db	2401	CTGCTGCTGTGCATCTTCAAAAGACTCGGGAGAAGACCTGGAGACCTCGGACCCTGTGTGACTTCATCTCA	2468
Qy	2461	GGGAGCAGAGATGCCCTTGAGACCCCATCTTAGATCTCAGAGAATTGAACCTTTGAAGCTGT	2520
Db	2461	GGGAGCAGAGATGCCCTTGAGACCCCATCTTAGATCTCAGAGAATTGAACCTTTGAAGCTGT	2520
Qy	2521	CCTAGTACCCAGATGTGTGATGTAGATNGCTCTGTGTTTTCTCAGGCCAAGGGACCTTAGAATGTGC	2580
Db	2521	CCTAGTACCCAGATGTGTGATGTAGATNGCTCTGTGTTTTCTCAGGCCAAGGGACCTTAGAATGTGC	2580
Qy	2581	TGACTTAATTAATTTTTTTTGTGATCTCTCACTCTGTGTTTTGGTTTTGGTTTTGGTTTTGG	2644
Db	2581	TGACTTAATTAATTTTTTTTGTGATCTCTCACTCTGTGTTTTGGTTTTGGTTTTGGTTTTGG	2644
Qy	2641	TTTTTGTTTTTTAAAGTGAATTTTGTCTGCTTTCAATTAATGTGAATCTGTGTTCTGGGAAAC	2700
Db	2641	TTTTTGTTTTTTAAAGTGAATTTTGTCTGCTTTCAATTAATGTGAATCTGTGTTCTGGGAAAC	2700
Qy	2701	TTCACCTGTGCCACTGAAGTTTATGTATACAGAGAATATTTGGCAATGATGTCCCTCTATTC	2766
Db	2701	TTCACCTGTGCCACTGAAGTTTATGTATACAGAGAATATTTGGCAATGATGTCCCTCTATTC	2766
Qy	2761	AAGGGGGGGTGGGGGGCGTTTTTCAAATGTATGTCTTGAGCACTGTCTGATTTGATGTCTCCA	2820
Db	2761	AAGGGGGGGTGGGGGGCGTTTTTCAAATGTATGTCTTGAGCACTGTCTGATTTGATGTCTCCA	2820
Qy	2821	GTCCCTTTCACACCCCAAGGCTGGCCACCTCCCTCATCTTATCTGTGTGGCCAAAAAAA	2880
Db	2821	GTCCCTTTCACACCCCAAGGCTGGCCACCTCCCTCATCTTATCTGTGTGGCCAAAAAAA	2880
Qy	2881	AAAAAAAAAAAAAAAAAAAAA 2902	
Db	2881	AAAAAAAAAAAAAAAAAAAAA 2902	
RESULT 3			
BC046833			
LOCUS			
DEFINITION	Mus musculus RIKEN cDNA 1200013B22 gene, mRNA (cDNA clone MGC:61244		
ACCESSION	IMAGE:6816569), complete cds.		
VERSION	BC046833		
KEYWORDS	BC046833.1 GI:29145027		
SOURCE	MGC.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 2938)		
	Straussberg,R.L., Peingold,E.A., Grouse,L.H., Derge,J.G.,		
	Klausner,R.D., Collins,F.S., Wagner,L.J., Shenmen,C.M., Schuler,G.D.,		
	Altshuler,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,		
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heile,F.,		
	Datchenko,L., Marinsina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
	Scapleton,M., Soares,W.B., Bonaldo,M.F., Casavant,T.L.,		
	Scheetz,T.E., Brownstein,M.U., Ussid,T.B., Toshayuki,S.,		
	Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,		
	Abrahamson,K.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,		
	Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,		
	Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,		
	Villalón,D.R., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
	Fahney,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,		
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		
	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,		
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
	Butterfield,Y.S., Krzywinski,M.T., Skalska,U., Smatius,D.E.,		
	Schmechel,A., Schein,J.E., Jones,S.J., and Marra,M.A.		
	Generation and initial analysis of more than 15,000 full-length		
	human and mouse cDNA sequences		
TITLE	JOURNAL MEDLINE PUBMED REFERENCE		
	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
	23388257		
	12477932		
	2 (bases 1 to 2938)		
AUTHORS	Straussberg,R.		

TITLE Direct Submission
JOURNAL Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kuceba, T., Lebeck, M., Melo, A., Schaefer, K., Schaefer, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

FEATURES
SOURCE
1. 2938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:61244 IMAGE:6816569"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NTH BMAP_FMO"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1. 2938
/gene="1200013B22Rik"
/note="synonym: Snack"
/db_xref="locusID:74137"
/db_xref="MGI:1921387"
81. 2000
/gene="1200013B22Rik"
/codon_start=1
/product="1200013B22Rik protein"
/protein_id="AAH4683.1"
/db_xref="GI:29145028"
/db_xref="locusID:74137"
/db_xref="MGI:1921387"
/translation="MESVALIQRPSPASALASBSARPLADGLIKSPKLMKKOAV
KRHHKNIHRVERFLTGKTYGVKAKBSGRLVAIKIKDKIKEDODLIHR
REIIMSLNPHITAIHEVGRSLVTVFENSRTIVMEYASGDLVDYDISEPRIS
ERDARHPFQIVSALHYCHONGIVARDKLENIILLDANNTKIDPGLSNLYHKFL
QTFCSPLVPIQVBEVNGKPYGPEVDWSLGLVLIYIVHGTMPDQDHTLVQISN
GAYREPKSPDACGLIRMLMVPTRATLEDAVSHMMVMNGYTGEOBALVEGQ
PSGDFGRASAMADMLRBSRPLNENGAKVCSFKOHVPGSGSTVGLRSHKIKSRK
NDMANONTGDPARTSSRPGKSLKLPGLIKKRSSTSSGVTGQDPRQIRVPTPBCQ
PVAVSLIPKRGITLKSRKQBSGYSSBPSBSGLLDASVFSGDEVKSKFSQASG
LLLRKGLIKNGFSRTRALEGTPSTFGSLDQLASSHPARPSRPGASVEDSILSS
ESPQDLPERLPTPLRGCVSVNLRGLLEQPPBEGIKRWQBSIGDSCEFLTDCQEV
TAAAYROALGICISKEYS"
249. 1025
/gene="1200013B22Rik"
/note="S_TKC; Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"
/db_xref="CDD:cd00180"

Query Match 97.2%; Score 2820.4; DB 10; Length 2938;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2855; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

43 GCGGCTGACCTCTGAGCCCGGCTCAGCGCCGCTGCTACTGCTGCTCCGACCCACTTC 102
|||||
Db 1 GCGGCTGACCTCTGAGCCCGGCTCAGCGCCGCTGCTACTGCTGCTCCGACCCACTTC 60
103 CACCTCGGGGCCCCGACCATGAGAGTGGGAGCTTACTCAGAGCCGACGAGCTC 162
61 CACTTCGGGTCGCCGACCATGAGAGTGGGAGCTTACTCAGAGCCGACGAGCTC 120
163 CCTGCGGCTCCGCCCTGCGCTCGAGAGCGCCGCGCTGCGAGCGGCTCATCAAGT 222
121 CCTGCGGCTCCGCCCTGCGCTCGAGAGCGCCGCGCTGCGAGCGGCTCATCAAGT 180
223 CGCTTAACTCTGATGAGAGAGAGCGGCTGAGAGCGGACCATCAAAACAACCTGC 282
181 CGCTTAACTCTGATGAGAGAGAGCGGCTGAGAGCGGACCATCAAAACAACCTGC 240
283 GGCACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGACCTAGGGAGGTGAAGAG 342
241 GGCACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGACCTAGGGAGGTGAAGAG 300
343 CACGAGAGAGCTCGGGCGTCTGGTGGCCATCAAGTCCATCAGAGAAAGACAAATCAA 402
301 CACGAGAGAGCTCGGGCGTCTGGTGGCCATCAAGTCCATCAGAGAAAGACAAATCAA 360
403 ATGAGCAGAGTCTGCTGACCATAGAGAGAGAGATTGAGATCATGTCTTCAACAACC 462
361 ATGAGCAGAGTCTGCTGACCATAGAGAGAGATTGAGATCATGTCTTCAACAACC 420
463 CCCACATCATGTCATCATCAAA-----GTGTTGAGATA 498
421 CCCACATCATGTCATCATCAAAAGTGGAGATCAAGTGTGATGATGATGAGATA 480
499 GCAGCAGAGTGTGATGTCATGAGAGTATGCGACCGAGGAGATCTGATGATGATCA 558
481 GCAGCAGAGTGTGATGTCATGAGAGTATGCGACCGAGGAGATCTGATGATGATCA 540
559 GTGAGCGGCAACGAGCTGAGTGAAGCGGAGCGGACGAGCATTTCTTCGACAGATCGTGTG 618
541 GTGAGCGGCAACGAGCTGAGTGAAGCGGAGCGGACGAGCATTTCTTCGACAGATCGTGTG 600
619 CCTGCACTACTGCAACGAGAGCGGATCGTTCAACGAGATCTCAAGTGTGAGAAACATCC 678
601 CCTGCACTACTGCAACGAGAGCGGATCGTTCAACGAGATCTCAAGTGTGAGAAACATCC 660
679 TTCTAGATGCAATGAGAAACATCAAGATGCTGCTTGGCTGCCAAGCTGTACACA 738
661 TTCTAGATGCAATGAGAAACATCAAGATGCTGCTTGGCTGCCAAGCTGTACACA 720
739 AAGCAAGTCTCTCAGACGTTCTGTGGAGACCTCTCTACGCTCGCTGAGATGATCA 798
721 AAGCAAGTCTCTCAGACGTTCTGTGGAGACCTCTCTACGCTCGCTGAGATGATCA 780
799 ACGGAAAGCCTTATGTGGGCGCCAGAGTGAACAGCTGCTCTGGGCGTTTCTCTGTACA 858
781 ACGGAAAGCCTTATGTGGGCGCCAGAGTGAACAGCTGCTCTGGGCGTTTCTCTGTACA 840
859 TCCGTGGTGCATGAGACCATGCGCTTTCAGCGGCGAGATCAAAACATGCTGAAGCAA 918
841 TCCGTGGTGCATGAGACCATGCGCTTTCAGCGGCGAGATCAAAACATGCTGAAGCAA 900
919 TCAGTAAACGAGGCTTACCGTGAAGCGGCGCAAGCGCTCGAGTCTGTGGCTGATCCGCT 978
901 TCAGTAAACGAGGCTTACCGTGAAGCGGCGCAAGCGCTCGAGTCTGTGGCTGATCCGCT 960
979 GGCCTTTAATGATGAACCCCAACCGCTGGGCGCAACATGAGAGATGAGCATGATGGT 1038
961 GGCCTTTAATGATGAACCCCAACCGCTGGGCGCAACATGAGAGATGAGCATGATGGT 1020

QY 1039 GGGTCACTGGGGTTTACACCAACCGAGTCGGGGAAACAGAAACCTCTGCTGAGGGTGGC 1038
Db 1021 GGGTCACTGGGGTTTACACCAACCGAGTCGGGGAAACAGAAACCTCTGCTGAGGGTGGC 1080
QY 1099 ACCCTAGTGGTGACTTTGGCCGGGCTTCATGCGGAGTGGTACGTGCTCTCGGCC 1158
Db 1081 ACCCTAGTGGTGACTTTGGCCGGGCTTCATGCGGAGTGGTACGTGCTCTCGGCC 1140
QY 1159 CCTCTGAGAGATGAGCCCAAGGTGTGACCTTTCAAGCAGACGTCGCGGAGGTG 1218
Db 1141 CCTCTGAGAGATGAGCCCAAGGTGTGACCTTTCAAGCAGACGTCGCGGAGGTG 1200
QY 1219 GAAGCAGTGAACCTGGGCTGAGAGGAGCAATTCCTTAAGAATCCCGAAAGAGATG 1278
Db 1201 GAAGCAGTGAACCTGGGCTGAGAGGAGCAATTCCTTAAGAATCCCGAAAGAGATG 1260
QY 1279 ACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGAGATCTCTTCTGACCTGGCAGA 1338
Db 1261 ACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGAGATCTCTTCTGACCTGGCAGA 1320
QY 1339 GCAGCCTTAAGCTTCGAAAGGCAATTCAGAAAGTCTCTACCTGTCAGGGAGG 1398
Db 1321 GCAGCCTTAAGCTTCGAAAGGCAATTCAGAAAGTCTCTACCTGTCAGGGAGG 1380
QY 1399 TACAGAGAGACCTTCAGAGAACTCAGACCGGTGCTGATCTCAGGGAGCTCTG 1458
Db 1381 TACAGAGAGACCTTCAGAGAACTCAGACCGGTGCTGATCTCAGGGAGCTCTG 1440
QY 1459 CTGATCTCTGCTCCCAAGAAAGCATCTTAAGAAGTCTCAGACGCTGATCTGGT 1518
Db 1441 CTGATCTCTGCTCCCAAGAAAGCATCTTAAGAAGTCTCAGACGCTGATCTGGT 1500
QY 1519 ACTACTCTCTCCAGAGCCCGAGAGTCTGGGAACTCTTAAGACGCCAGTATGTTG 1578
Db 1501 ACTACTCTCTCCAGAGCCCGAGAGTCTGGGAACTCTTAAGACGCCAGTATGTTG 1560
QY 1579 TGAAGTGGGAGCCCGTGGAGAGAAAGTCTCAGACGCTTCAGGGCTCTCTCAACGCA 1638
Db 1561 TGAAGTGGGAGCCCGTGGAGAGAAAGTCTCAGACGCTTCAGGGCTCTCTCAACGCA 1620
QY 1639 AGGGCATCTCAAACTCAATGAGCAAGTTCCTCCGCAAGCTTGAAGGCACTACCCCTA 1698
Db 1621 AGGGCATCTCAAACTCAATGAGCAAGTTCCTCCGCAAGCTTGAAGGCACTACCCCTA 1680
QY 1699 GCACCTTTGGCTCCCTGAGACCAACTGGCTCTCCATCTCAGACCCCGGCCAGCC 1758
Db 1681 GCACCTTTGGCTCCCTGAGACCAACTGGCTCTCCATCTCAGACCCCGGCCAGCC 1740
QY 1759 CTTCAAGGGGCTGTGAGTGAAGAGCAATCTGTCTCTCGAGTCTCTTGAACCAATTGACT 1818
Db 1741 CTTCAAGGGGCTGTGAGTGAAGAGCAATCTGTCTCTCGAGTCTCTTGAACCAATTGACT 1800
QY 1819 TGCCTGAACGTCCTCCGGAACCCCACTGAGAGGGCTGTGCTGTGAGCAACCTGAGGG 1878
Db 1801 TGCCTGAACGTCCTCCGGAACCCCACTGAGAGGGCTGTGTCTGTGAGCAACCTGAGGG 1860
QY 1879 GGGTTCAGAGCCTCCCTCAGAAAGTCTGAGCGATGGTGGCAGAAATCTTTGGGGATA 1938
Db 1861 GGGTTCAGAGCCTCCCTCAGAAAGTCTGAGCGATGGTGGCAGAAATCTTTGGGGATA 1920
QY 1939 GCTGCTTTTCTCTGACAGACTGCCAAGAGTGACTGACAGCTTACAGACAGACCTTAGAA 1998
Db 1921 GCTGCTTTTCTCTGACAGACTGCCAAGAGTGACTGACAGCTTACAGACAGACCTTAGAA 1980
QY 1999 TCTGCTCAAAAGCTCAGCTGAGAGAGAGAGATGAGCTTATGAGGGTGGCTCTAGA 2058
Db 1981 TCTGCTCAAAAGCTCAGCTGAGAGAGAGAGATGAGCTTATGAGGGTGGCTCTAGA 2040
QY 2059 GGGTTTCAGAGAACCTGGTGGATCTCTCCAGTGAATAGATACATCAAGGGCTCT 2118
Db 2041 GGGTTTCAGAGAACCTGGTGGATCTCTCCAGTGAATAGATACATCAAGGGCTCT 2100
QY 2119 ACCTTCAGAGCTGACTGAACTGAAAGATGAGAAATCCGATTTGATGTGAAAGAT 2178

Db 2101 ACCTTCAGAGCTGACTGAACTGAAAGATGAGAAATGCAATGATGTGAAAGAT 2160
QY 2179 GGGAAACCTTGCTGCCGAGTGTATAGTGGGTGGCTGAAGGTGCTACCTCTTGT 2238
Db 2161 GGGAAACCTTGCTGCCGAGTGTATAGTGGGTGGCTGAAGGTGCTACCTCTTGT 2220
QY 2239 GCCATGAGTGCACCCATGACATTTCCACCTGTTCTGTGGCTGACCTTCATAGT 2298
Db 2221 GCCATGAGTGCACCCATGACATTTCCACCTGTTCTGTGGCTGACCTTCATAGT 2280
QY 2299 TTCTGTTCCATCAACACAGAGTTGAACCTGATCTTCTGGAGTATGTGTGTG 2358
Db 2281 TTCTGTTCCATCAACACAGAGTTGAACCTGATCTTCTGGAGTATGTGTGTG 2340
QY 2359 ACTGCAATTTTGAAGAGAAACAAGCTCTGGTTTCATCTCTGCTGTGATCTCA 2418
Db 2241 ACTGCAATTTTGAAGAGAAACAAGCTCTGGTTTCATCTCTGCTGTGATCTCA 2400
QY 2419 AAGACTGGGAGACTCGACCGCTGTTTGAATTCATCTCAAGGAGACAGATGCCCTG 2478
Db 2401 AAGACTGGGAGACTCGACCGCTGTTTGAATTCATCTCAAGGAGACAGATGCCCTG 2460
QY 2479 GACCCCATCTTATATCTCAGAGACTTGAACCTTGAAGCTGTCTTATACCCAGATGG 2538
Db 2461 GACCCCATCTTATATCTCAGAGACTTGAACCTTGAAGCTGTCTTATACCCAGATGG 2520
QY 2539 ATGATGCTCTGTTTTCAGAGCAACGAGACTGAATGCTGACTTATTTT 2598
Db 2521 ATGATGCTCTGTTTTCAGAGCAACGAGACTGAATGCTGACTTATTTT 2580
QY 2599 GTGATCTCACTCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2658
Db 2581 GTGATCTCACTCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2640
QY 2659 TTTTGTCTTCAATTAATGTAATGCTGTCTGGGGAACCTCACTGAGCACTAAG 2718
Db 2641 TTTTGTCTTCAATTAATGTAATGCTGTCTGGGGAACCTCACTGAGCACTAAG 2700
QY 2719 TTTATGACAGAGAAATTTGGCAATGATGCTCTTATTTCAAGGGGGTGGGGGCTT 2778
Db 2701 TTTATGACAGAGAAATTTGGCAATGATGCTCTTATTTCAAGGGGGTGGGGGCTT 2760
QY 2779 TTTCAATGTATGCTTGAACACTGTCTGATTTGATGCTTCAGTCCCTTCAACCCAGG 2838
Db 2761 TTTCAATGTATGCTTGAACACTGTCTGATTTGATGCTTCAGTCCCTTCAACCCAGG 2820
QY 2839 CTGGCCACCTCTCCCTCATCTTCATCTGTGGCCAAAAAATTTTTTTTTT 2898
Db 2821 CTGGCCACCTCTCCCTCATCTTCATCTGTGGCCAAAAAATTTTTTTTTT 2880

RESULT 4
BC033302 2917 bp mRNA linear ROD 04-NOV-2003
LOCUS BC033302
DEFINITION Mus musculus RIKEN cDNA 1200013b22 gene, mRNA (cDNA clone MGC:30598
IMAGE:3663595), complete cds.
ACCESSION BC033302
VERSION BC033302.1 GI:23271085
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 2917)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Dietzen, R.J., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uebachs, T.B., Tomshyuk, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLES
JOURNAL

REMARK
COMMENT

Abrahamson, R.D., Mullaly, S.J., Boesak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gamaral, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.R., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Ketterman, M., Madan, A.C., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Maxam, M.A.
human and mouse cDNA sequences of more than 15,000 full-length
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 2917)
Strausberg, R.
Direct Submission
Submitted (26-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
DNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gamaral, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kowit, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Manavati,
A.N., Gibbs, R.A.

FEATURES
SOURCE

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 40 Row: d Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27229181.
Location/Qualifiers

1. .2917
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="MGC:30598 IMAGE:3663595"
/tissue_type="Mammary tumor metastasized to lung."
MTV-LTR/mtl model. Expression driven by an MTV-LTR
enhancer."
/clone_id="NCI CGAP Lu30"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2917
/gene="1200013B22R1k"
/note="Synonym: Shark"
/db_xref="LocusID:74137"
/db_xref="MGI:1921387"
81. .1976
/gene="1200013B22R1k"
/codon_start=1
/product="1200013B22R1k protein"
/protein_id="AAH3302.1"
/db_xref="GI:23271086"
/db_xref="LocusID:74137"
/db_xref="MGI:1921387"
/translation="MESVALLOPSPQASALASASAPLADGLIKSPIMKKOAV
KHHHNNLRHYELETGKGTGKPKASGRVLAIRPKDKPODLHTR
REIFNSLNPPIIAIHEVFNSSKIVIMETASRDLDYISERPLSRDRHFF
ROIVALHCHONGIIVARDKLENIILDAANGNLIKADPGSNLYHKGKLTQTCGSL
YASPEIVNGRPYGPEDVSWSLGLVLYLVHGTPFDGQHKTLVKQISNGAYREPK

misc_feature
/gene="1200013B22R1k"
/note="S TKc. Region: Serine/threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"

ORIGIN

Query Match 95.2%; Score 2763.8; DB 10; Length 2917;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2808; Conservative 0; Mismatches 37; Indels 2; Gaps 2;

55 TCTGAGCCCGCGGCTCAGCGCGCTCTACTGCTGCCGACCACTCACTCGCGCTC 114
13 TCTGAGCCCGCGGCTCAGCGCGCTCTACTGCTGCCGACCACTCACTCGCGCTC 72
115 CCGGACACATGAGTGGTGGCTTACTCCAGCGCCCGAGCCGAGCTCCCTGGCTCCG 174
73 CCGGACACATGAGTGGTGGCTTACTCCAGCGCCCGAGCCGAGCTCCCTGGCTCCG 132
175 CCGTGGCTCGGAGAGCGCCCGGCTGGCGGAGCGGCTCATGAGTGGCTTAACTC 234
133 CCGTGGCTCGGAGAGCGCCCGGCTGGCGGAGCGGCTCATGAGTGGCTTAACTC 192
235 TGATGAAGGAGGAGCGCGGTGAAGGCGGACCATCAAAACAACCTCGGACCGCTACG 294
193 TGATGAAGGAGGAGCGCGGTGAAGGCGGACCATCAAAACAACCTCGGACCGCTACG 252
295 AGTTCTGAGACGCTGGCGGACGCTTACCGGAAAGTGAAGAGGACGACGAGAGCT 354
253 AGTTCTGAGACGCTGGCGGACGCTTACCGGAAAGTGAAGAGGACGACGAGAGCT 312
355 CGGGCGCTCGTGGCGGACGCTTACCGGAAAGTGAAGAGGACGACGAGAGCT 414
313 CGGGCGCTCGTGGCGGACGCTTACCGGAAAGTGAAGAGGACGACGAGAGCT 372
415 TGCTGACATACGAGGAGGAGTGAATGATGCTTCACTCAACCACTCCGACATCATTTG 474
373 TGCTGACATACGAGGAGGAGTGAATGATGCTTCACTCAACCACTCCGACATCATTTG 432
475 CCATCGATGAGTGGTGAATGAGGAGGAGTGAATGATGCTTCACTCAACCACTCCGAC 534
433 CCATCGATGAGTGGTGAATGAGGAGGAGTGAATGATGCTTCACTCAACCACTCCGAC 492
535 GAGGCGATCTGTATGATTAATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 594
493 GAGGCGATCTGTATGATTAATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 552
595 ATTTCTTCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
553 ATTTCTTCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
655 GAGATCTCAAGCTGAGAAACATCTTCTAGATGCAATGAGAAACATCAAGATTGCTGACT 714
613 GAGATCTCAAGCTGAGAAACATCTTCTAGATGCAATGAGAAACATCAAGATTGCTGACT 672
715 TTGGCCCTTCGACCTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 774
673 TTGGCCCTTCGACCTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 732
775 TCTACGCGCTGCTGAGATAGTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 834

Db 733 TCTACGCTCCCTGGAATGATCAACGGGAGCCCTATGTGGCCCAAGGTGACACT 792
Qy 835 GGTCTGCGGCGTTCTCTGTACATCCTGTGATGCGACCATGCCCCCTTTGACGGGACG 894
Db 793 GGTCTGCGGCGTTCTCTGTACATCCTGTGATGCGACCATGCCCCCTTTGACGGGACG 852
Qy 895 ATCTATAAACACTGTGTAAGAAATCAATACGGGGCTTACCGGTGAGCCGCCCAAGCCGT 954
Db 853 ATCTATAAACACTGTGTAAGAAATCAATACGGGGCTTACCGGTGAGCCGCCCAAGCCGT 912
Qy 955 CCGATGCGCTGTGCGCTGATCCGGTGGCTGTTAATGTGTAAACCCACCCCGTGGGCAAC 1014
Db 913 CCGATGCGCTGTGCGCTGATCCGGTGGCTGTTAATGTGTAAACCCACCCCGTGGGCAAC 972
Qy 1015 TGGAGATGTAGCCAGTCATTTGTGTGCTCAACTGGGGTTTACCAACCGGAGTGGGGAA 1074
Db 973 TGGAGATGTAGCCAGTCATTTGTGTGCTCAACTGGGGTTTACCAACCGGAGTGGGGAA 1032
Qy 1075 AGGAAGCCCTGCGGAGGAGGAGGAGCCCTAGTGTGACTTTGGCCGGGCTTCATGTGCG 1134
Db 1033 AGGAAGCCCTGCGGAGGAGGAGGAGCCCTAGTGTGACTTTGGCCGGGCTTCATGTGCG 1092
Qy 1135 ACTGTTACGTCGCTCTCGGCCCCCTCTGTAGAAATGAGCCAGGTGTGACGTTCT 1194
Db 1093 ACTGTTACGTCGCTCTCGGCCCCCTCTGTAGAAATGAGCCAGGTGTGACGTTCT 1152
Qy 1195 TCAAGCAGCAGTGCCTGGGAGGTGAAAGCACTGTACCTGGGCTGAGCGGCAATTC 1254
Db 1153 TCAAGCAGCAGTGCCTGGGAGGTGAAAGCACTGTACCTGGGCTGAGCGGCAATTC 1212
Qy 1255 TTTAAGAGTCCCGGAAAGAAATGACATGTGCTCAAAATCTGCAAGGTGACCCGGCTAG 1314
Db 1213 TTTAAGAGTCCCGGAAAGAAATGACATGTGCTCAAAATCTGCAAGGTGACCCGGCTAG 1272
Qy 1315 ATACCTCTTCTCGGCTGSCAAAGAGCAGCCCTTAAGCTTCGAAAGGCAATTCGAA 1374
Db 1273 ATACCTCTTCTCGGCTGSCAAAGAGCAGCCCTTAAGCTTCGAAAGGCAATTCGAA 1332
Qy 1375 AGTCTCTACCTCTGTCAGGGAGGTAGACAGAGACCCCTCAGAACTCAGACCGGTGCTG 1434
Db 1333 AGTCTCTACCTCTTCAAGGGAGGTAGACAGAGACCCCTCAGAACTCAGACCGGTGCTG 1392
Qy 1435 ATACTCCAGGAGCAGCCTGTCTCTGTATCTCTGTCTCCAGGAAAGGCACTCTTAA 1494
Db 1393 ATACTCCAGGAGCAGCCTGTCTCTGTATCTCTGTCTCCAGGAAAGGCACTCTTAA 1452
Qy 1495 AGTCTCAGACGCGTAATCTGTACTACTCTCTCCAGAGCCAGGAGCTGGGAAAC 1554
Db 1453 AGTCTCAGACGCGTAATCTGTACTACTCTCTCCAGAGCCAGGAGCTGGGAAAC 1512
Qy 1555 TCTTGAACGCGCAGT 1614
Db 1513 TCTTGAACGCGCAGT 1572
Qy 1615 CTTCAGGGGCTCTCTCTCAACGCAAGGCAATTCGAAATCTCAATGCGAAGTTCTCCG 1674
Db 1573 CTTCAGGGGCTCTCTCTCAACGCAAGGCAATTCGAAATCTCAATGCGAAGTTCTCCG 1632
Qy 1675 CAGCCTTAAGAGGCACTACCCCTTAGCACTTTGGCTCCCTGAGCAACATGCGCTCC 1734
Db 1633 CAGCCTTAAGAGGCACTACCCCTTAGCACTTTGGCTCCCTGAGCAACATGCGCTCC 1692
Qy 1735 ATCTGAGGCGCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGAGACATCTGTCT 1794
Db 1693 ATCTGAGGCGCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGAGACATCTGTCT 1752
Qy 1795 CCGAGTCTTTTGAACCAATTGACTTGTGCTTCCGAAACCTTCCGAAACCTCAGGGGCT 1854
Db 1753 CCGAGTCTTTTGAACCAATTGACTTGTGCTTCCGAAACCTTCCGAAACCTCAGGGGCT 1812
Qy 1855 GTGTGTGTGTGAACAACCTAGGGGGCTTGAAGAGCTCTCCCTGAAAGTCTGAAGGAT 1914

Db 1813 GTGTGTGTGTGAACAACCTAGGGGGCTTGAAGAGCTCCCTCAGAAAGTCTGAAGGAT 1872
Qy 1915 GGTGAGAGAAATCTTGGGGGATAGCTGTTTTCTGTAGACACTGCAAGGTGA 1974
Db 1873 GGTGAGAGAAATCTTGGGGGATAGCTGTTTTCTGTAGACACTGCAAGGTGA 1932
Qy 1975 CAGCCTTAAGAGGCGCTTGAATCTGTCTCAAAAGCTCACTGAGGAAGGAGATGTGTC 2034
Db 1933 CAGCCTTAAGAGGCGCTTGAAGATCTGTCTGAAGCTCACTGAGGAAGGAGATGTGTC 1992
Qy 2035 CTTAGTATGGGGTATAGGCTCTGAGAGGGTGTGCAAGAGAAACCTTGGGTGAGTTCT 2094
Db 1993 CTTAGTATGGGGTATAGGCTCTGAGAGGGTGTGCAAGAG-ACCTTGGGTGAGTTCT 2051
Qy 2095 TGAATAGATGATCAAGAGGCTCTACCTGTGAGCCTGACTGAACCTGAAAGATGAGA 2154
Db 2052 TGAATAGATGATCAAGAGGCTCTACCTGTGAGCCTGACTGAACCTGAAAGATGAGA 2111
Qy 2155 AATCGCATGTATGTGAAAGAAATGGGAACCTTGTGCTGCCCGAGTGTATAGTGGGTG 2214
Db 2112 AATCGCATGTATGTGAAAGAAATGGGAACCTTGTGCTGCCCGAGTGTATAGTGGGTG 2171
Qy 2215 CTTGAAGGTCCTACCTCTTGTGACATGATGTCACCCATGACATTTCCACCCGT 2274
Db 2172 CTTGAAGGTCCTACCTCTTGTGACATGATGTCACCCATGACATTTCCCGCCTGT 2231
Qy 2275 CTCTGGCTGACCTTCAATTAAGTTCGTTTCATCAACACCAAGGTTAGAACCTGA 2334
Db 2232 CTCTGGCTGACCTTCAATTAAGTTCGTTTCATCAACACCAAGGTTAGAACCTGA 2291
Qy 2335 CTTCCTGGAGGTATGTATGTGATGTCGCAATTTATAGAGAGAAACACCTCTGTT 2394
Db 2292 CTTCCTGGAGGTATGTATGTGATGTCGCAATTTATAGAGAGAAACACCTCTGTT 2351
Qy 2395 CCAATCTGCTGCTGTCATCTCAAGAACCTGGGAGAGCTGGAGCCGCTGTTGACTCA 2454
Db 2352 CCAATCTGCTGCTGTCATCTCAAGAACCTGGGAGAGCTGGAGCCGCTGTTGACTCA 2411
Qy 2455 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTGAATCTCAGAGACTTGAAC 2514
Db 2412 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTGAATCTCAGAGACTTGAAC 2471
Qy 2515 GCTGTCTTGTATCCAGATGTGATGTGATGTCTGTCTCAGGCCAAGGAGCTTGA 2574
Db 2472 GCTGTCTTGTATCCAGATGTGATGTGATGTCTGTCTCAGGCCAAGGAGCTTGA 2531
Qy 2575 ATGTCGATCTTAATTTATTTTGTGATCTGACCTCTGTTTTGTTTTGTTTTGTT 2634
Db 2532 ATGTCGATCTTAATTTATTTTGTGATCTGACCTCTGTTTTGTTTTGTTTTGTT 2591
Qy 2635 TGT 2693
Db 2592 TGT 2651
Qy 2694 GGGGAATCTCACTGTGCTGATGAAGTTTATGTACAGAGAGTATTTGGCAATGTGCC 2753
Db 2652 GGGGAATCTCCCTGTGCTGATGAAGTTTATGTACAGAGAGTATTTGGCAATGTGCC 2711
Qy 2754 TCTATTCAAGGGGGGTGGGGGCTTTTCAAAAGTATGTCTGAGCACTGTGGAATGA 2813
Db 2712 TTTATTCAAGGGGGGTGGGGGCTTTTCAAAAGTATGTCTGAGCACTGTGGAATGA 2771
Qy 2814 GTCTCAGTCCCTTCAACCAAGGTGAGCACTTCCCTCATCTTCAATGTGTGCAAA 2873
Db 2772 GTTTGAGTCCCTTCAACCAAGGTGAGGCCCCCTCCCTCATTTCTTTGTGTGCAAA 2831
Qy 2874 AAAAAAAAAAAAAAAAAAAAAA 2900
Db 2832 AAAAAAAAAAAAAAAAAAAAAA 2858
RESULT 5
BC081899

LOCUS BC081899 2900 bp mRNA linear ROD 02-SEP-2004
DEFINITION Rattus norvegicus cDNA clone MGC:93817 IMAGE:7110518, complete cds.
ACCESSION BC081899
VERSION BC081899.1 GI:51859485
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2900)
Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Uebin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shcherbenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smailus, D.E.,
Schmechel, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2900)
Director MGC Project.
Direct Submission
Submitted (01-SEP-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
Series: IRAC Plate: 182 Row: a Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2900
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93817 IMAGE:7110518"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_id="NIH_MGC_235"
/lab_host="DH10B"
/note="Vector: pExpress1"
73..1965
/codon_start=1
/product="unknown (protein for MGC:93817)"
/protein_id="AAH81899.1"

/db_xref="GI:51859486"
/translation="MESVALHRRGNLAPSAALATBSARPLADRLIKSPYPLMKQAV
KRHHKHNILHRYEFLTEKGTGYKPKKRSRGLVAIKSRKDKIKQEDILHTR
REIEMSLNHPHIAHVEFENSKVIWMEVASRDLVDYISERPLRDLRHF
ROIYVALYCHONGIYHRDLKENILLDAGNLIKADPGSLNYHKGKPIOTPCGSL
YASPELVNRPYVGPVDSVSLVLYIYVHGTMPPGQDHKLYKQISSGAYRBP
PSDACGLIRMLNVPYRATLIEDVASHMWNNGYSTRIEQLERKGRHPSGDSGA
SMADWLRSSRPILLENGAKVCSFKHVPVGSSTGRLPQHSIKSRKNDMAQTLQ
NDPAEOTSRRPGNSVLEPKGILKKKASPSSEGEVQEPDLRPVSNTPQVPAIPIL
PRKGIKKSRRESGYVSPSESGELDAGVFGVSGDPEQSPASGRILHRKGLI
KNGKFSRTLFGTARSTFGSLDOLASPHPTARASRPSGASVSDSLSSRFDLDLP
ERLPYPLRQCVSDVNDLRLEQFPSESLKRWQBSLSDSCPSLTDCEVTAAIRQALG
ICSKLS"

ORIGIN
Query Match 72.5%; Score 2102.6; DB 10; Length 2900;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2507; Conservative 0; Mismatches 259; Indels 109; Gaps 14;
51 GACCTTGAGCCCGCGGCTGACGCGGCTGCTACTGCTGCCGACCCACTTCACCTGCG 110
1 GACCTTGAGCTGCGGCTCTCCGCGGCTGCTGCTGCTGCTGCCGACCCCTCGGCTGCG 60
QY 111 GGTCCCGCACCATGAGTGTGCTTACTTCAAGCCCGGACCAAGCTCCTCGGCG 170
61 CGTCCCGCACCATGAGTGTGCTTACTTCAAGCCCGGACCAAGCTCCTCGGCG 120
QY 171 TCCGCGCTGCGCTGCGGAGACCGCCCGGCGGACCGGCTCATCAAGTGGCTTAA 230
121 TCCGCGCTGCGCTGCGGAGACCGCCCGGCGGACCGGCTCATCAAGTGGCTTAA 180
QY 231 CCTCTGATGAAGAAGACGAGCGGTGAAGCGGACCATCAACAACACTGCGGACCGG 290
181 CCTCTGATGAAGAAGACGAGCGGTGAAGCGGACCATCAACAACACTGAGCACCGG 240
QY 291 TACGAGTTCTTGAGAGCGCTGCGGCAAGGCACTTACGGAAGGTGAAGAGCGAG 350
241 TACGAGTTCTTGAGAGCGCTGCGGCAAGGCACTTACGGAAGGTGAAGAGCGAG 300
Db 351 AGCTCGGCGGCTGTGCTGCTGCTTCAATCCATCAGGAAGCAAAATCAAGATGAGAG 410
301 AGCTCGGCGGCTGTGCTGCTTCAATCCATCAGGAAGCAAAATCAAGATGAGAG 360
QY 411 GATCTGTGACATATACGAGGAGATGAGATGATGATCTTCACTCAACACCCCACTC 470
361 GATCTGTGACATATACGAGGAGATGAGATGATGATCTTCACTCAACACCCCACTC 420
QY 471 ATTGCGATTCATGAGTGTGAGAAATAGACGCAAGATTGTGATGTGATGATATGCC 530
421 ATTGCGATTCATGAGTGTGAGAAATAGACGCAAGATTGTGATGTGATGATATGCC 480
Db 531 AGCCGAGCGCATCTGTATGATTATCATATGATGAGCGGCGCATGAGTGAAGCGGAG 590
481 AGCCGAGCGCATCTGTATGATTATCATATGATGAGCGGCGCATGAGTGAAGCGGAG 540
QY 591 AGGATTTCTTCGACGAGATGCTGTGCGCTGACACTGACCAAGCGGATGCT 650
541 AGGATTTCTTCGACGAGATGCTGTGCGCTGACACTGACCAAGCGGATGCT 600
Db 651 CACCGAGATTCATGAGTGAAGAAACATCTTCTGATGCTCATGAGAAACATCAAGATTGCT 710
601 CACCGAGATTCATGAGTGAAGAAACATCTTCTGATGCTCATGAGTGAAGAAACATCAAGATTGCT 660
QY 711 GACTTTGGCTCTTCAACCTGTATACCAAGGCAAGTTCTTCAAGAGTTTGTGGAGAC 770
661 GATTTTGGCTCTTCAACCTGTATACCAAGGCAAGTTCTTCAAGAGTTTGTGGAGAC 720
Db 771 CCTCTCAAGCTGCGCTGAGATGATGATCAAGGGAAGCGCTATGAGGCGGACGAGAGTGAAC 830
721 CCTCTCAAGCTGCGCTGAGATGATGATCAAGGGAAGCGCTATGAGGCGGACGAGAGTGAAC 780
QY 831 AGCTGATCTGTGGCGTTCTCTGTATATCATCTGCTGATGAGGACATGCTTGTGACGGG 890

Db 781 AGCGGCTCTGCGGCGTTCTTCTGTACATCCGTGTCATGSCACCATGCCCTTTGAAGG 840
QY 891 CAGGATATATAAAGCTGTGAGACAAATCATGATACGCGGCTTACCGTAGCCGCCAAG 950
Db 841 CAGGATATATAAAGCTGTGAGACAAATCATGATACGCGGCTTACCGTAGCCGCCAAG 900
QY 951 CCGTCCGATGCTGTGCGCTGATCCGGTGGCTGTTAATGAGAACCCCGCGGCGG 1010
Db 901 CCGTCTGATGCTGTGCGCTGATCCGGTGGCTGTTAATGAGAACCCCGCGGCGG 960
QY 1011 ACACTGAGAGATGATGACAGTCAATGATGAGTCAACTGCGGCTTACACACCGAGTGG 1070
Db 961 ACTGAGAGATGATGACAGTCAATGATGAGTCAACTGCGGCTTACACACCGAGTGG 1020
QY 1071 GAACAGAGAGCCCTGCGCTGAGGCTGGGACCTGATGATGATCTTTGCGCGGCTTCAATG 1130
Db 1021 GAACAGAGAGCCCTGCGAGAGGCTGGGACCTGATGATGATCTTTGCGCGGCTTCAATG 1080
QY 1131 GCGGACTGTTAGTGTGCTCTGCGGCGGCTTCTGAGAGATGAGACCAAGTGTGAGC 1190
Db 1081 GCGGACTGTTAGTGTGCTCTGCGGCGGCTTCTGAGAGATGAGACCAAGTGTGAGC 1140
QY 1191 TTCTTCAAGAGACGCTGCGGAGTGTGAAAGACTGTACCTGCGGCTGAGCGGCAACAT 1250
Db 1141 TTCTTCAAGAGACGCTGCGGAGTGTGAAAGACGCGGACCGGCTGAGCGGCAACAT 1200
QY 1251 TCTCTTAAGAGTCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCGGCT 1310
Db 1201 TCTCTTAAGAGTCCGCAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCGGCT 1260
QY 1311 GAGGATACCTCTTCTGCGGCTGCGAGAGAGCTTAAAGCTTCCGAAAGGATCTCAAG 1370
Db 1261 GAGGATACCTCTTCTGCGGCTGCGAGAGAGCTTAAAGCTTCCGAAAGGATCTCAAG 1320
QY 1371 AAAAAGCTCTTACTCTGTCAGAGGAGTACAGAGAGACCTCAGAACTCAGACCGGCTG 1430
Db 1321 AAAAAGCTCTTCTGTCAGAGGAGTACAGAGAGGCTCCTCAGAACTCAGACCGGCTG 1380
QY 1431 CCGTATCTCAGAGGAGCTCTGCTGCTGCTGATCTCTGCTCCAGAGAAAGGATCTT 1490
Db 1381 TCCAAATCCCGAGGAGCTGCTGCTGCTGATCTCTGCTCCAGAGAAAGGATCTT 1440
QY 1491 AAGAAGCTGAGACGCGTGAATCTGTTACTACTCTCTCCAGAGGCCAGAGATCTGGG 1550
Db 1441 AAGAAGCTGAGACGCGTGAATCTGTTACTACTCTCTCCAGAGGCCAGAGATCTGGG 1500
QY 1551 GAACTCTTAGACGCGAGTGAATGTTGTGAGTGGGACCCCGAGAGAGAAATCTCCA 1610
Db 1501 GAACTCTTAGACGAGGTGAATGTTGTGAGTGGGACCCCGAGAGAGAAATCTCCA 1560
QY 1611 CAGGCTTACGAGGCTCTCTCTCCACCGCAAGGAGCTTCAAACTCAATGGCAAGTTCTCC 1670
Db 1561 CAGGCTTACGAGG---CGCTCTCAATCGCAAGGAGCTTCAAACTCAATGGCAAGTTCTCC 1617
QY 1671 CGACAGGCTTAAAGGACCTACCCCTTAAGACCTTTGGCTCCCTGAGACCAACTGGCTCC 1730
Db 1618 CGACAGGCTTAAAGGACCTACCCCTTAAGACCTTTGGCTCCCTGAGACCAACTGGCTCC 1677
QY 1731 TCCCATCTGACGCGCGGCGACGCGGCTCAGAGGCTGATGATGAGACAGCATCTG 1790
Db 1678 CTTCACTCTACAGCGCGGCGACGCGCTCTCGAGAGCTGATGATGAGACAGCATCTG 1737
QY 1791 TCTCTCGAGTCTTTTGAACAATGGAAGCTTGGCTGAAAGTCTTCCGAAACCCCACTAGG 1850
Db 1738 TCTCTCGAGTCTTTTGAACAATGGAAGCTTGGCTGAAAGTCTTCCGAAACCCCACTAGG 1797
QY 1851 GCGTGTCTCTGTGAGCAACTGAGGAGGCTTGAAGAGCTCTCCCTAGAAAGTCTGAAG 1910
Db 1798 GCGTGTCTCTGTGAGCAACTGAGGAGGCTTGAAGAGCTCTCCCTAGAAAGTCTGAAG 1857
QY 1911 CGATGCTGAGAGATCTTGGGAGATAGCTGCTTTCTCTGACAGACTGCGCAAGAGAGT 1970
Db 1858 CGATGCTGAGAGATCTTGGGAGATAGCTGCTTTCTCTGACAGACTGCGCAAGAGAGT 1917

QY 1971 ACTGACGCTTACAGACAGGCTTACAGATCTGCTCAAAAGTCACTGAGAGAGAGATG 2030
Db 1918 ACAGACGCTTACAGACAGGCTTACAGATCTGCTCAAAAGTCACTGAGAGAGAGATG 1977
QY 2031 GTGCGCTAGT-ATGGGCTAGGCTCTGAGA-GGGTTTGCAGAGAAACCTGGGTGATTC 2088
Db 1978 GTGCGCTAGTATGGGCTAGGCTCTGAGAGGGTTTGCAGAGAAACCTGGGTGATTC 2037
QY 2089 CTCACGTAATGATGATCAATCAAG---GGCTCTACGCTGACGCTGATCACTGAAC 2144
Db 2038 CCCAGGGAGATGATGATCAATCAAGAACTCTCTGCTGCTTCAAGCTGATGAAACCTGA 2097
QY 2145 AGATGAGAAATCGCAATGATGTAAGAAATGCG-----GACCTTGTCTGCCGAG 2198
Db 2098 GGGTGAAGAAATGAGAGATGTAAGAAAGTCACTGACAGAGTCTTCACTGCCAGAA 2157
QY 2199 TGTATATAGGGGT-----GGCTGAAGGTGCTACCTCTTTGGCCATGATGTCAC 2252
Db 2158 TGTTCAGCGGGGACACAGAGACTTGAAGAGTCTTCTCTTTATGC--TGAAGTCTAC 2215
QY 2253 CCATGACATTTCCACCTGCTCTGCTGACAC-----TTCACATAGTTCTGT 2304
Db 2216 CCAATGGATCTCC-CCTGCTCTGCGCAGTGTCAAGGCTGATCCACATAGTCTGT 2274
QY 2305 TTCAATCAACACAGGCTTGAACCTGACTTCTGCGAGGATATGTGATGATGCC 2364
Db 2275 TCCGATCGACACAGGCTTGAACCTGACTTCTGCGAGGATATGTGAGAACCTCG 2334
QY 2365 ATTATTTAGAGAGAAACAGCTCTGCTTCAATCTGCTGCTGATCTGAAGAC 2424
Db 2335 CTTATTTAAGAGAAACAGCTCTGCTTCAATCTGCTGCTGATCTGAAGAC 2394
QY 2425 TGGAAAGACTCGACCGCTGTTGACTTCACTCAAGGAGACCAATGCCCTGAACCC 2484
Db 2395 TAGAAAGACTC-AACGCTGTTCACTTCACTCAAGGAGAC----- 2435
QY 2485 ATCTTGAATCTCAGACCTTGAACCTTGAAGCTGTTCTAGTACCCAGATGATGAT 2544
Db 2436 -----CTCAGAGACCTGAGCTTGAAGCTGTTCTGATTAACAGACTATGATGAT 2486
QY 2545 -GCTGTTTCTCAGGCGAACCGGACCTTGAAGTGTGACTTATTTATTTT----- 2596
Db 2487 ATGCTGTTTCTCAGGCGAACCGGACCTTGAAGTGTGACTTATTTATTTT----- 2546
QY 2597 -----TTGATCTCACTTCTGTTTGTGT 2623
Db 2547 TCACTTCTGCTCTGTTCTGTTTCTGTTTCTGTTTGTGTTTGTGTTGTTTCT 2606
QY 2624 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGAAGTGAATTTGCTCTTCAATTAATGAT 2683
Db 2607 TGTGTTTGTGTTTGTGTTTGTGTTTGAAGTGAATTTGCTCTTCCGTTAATGAT 2666
QY 2684 GCTGTTTCTGAGGAACTCACTGTCGCACTGAAGTTAATGATCAAGAGATTTGGCA 2743
Db 2667 GCTGTTTCTGAGGAACTCACTGTCATTTGAAGTGTGATCAAGAGATTTGGCA 2726
QY 2744 ATGATGCTCTTATTAAGAGGAGTGGGAGCTTTTCAATGATATGTTGAGACAGT 2803
Db 2727 GTGATCTCTTAT---GGGGGTGCTTTTCAAGTATGTTTGAAGACTG 2778
QY 2804 TCTGATGATGATCTCAGTCTTCAACCCAGAGGCTGCGACCTTCCCTCATCT 2858
Db 2779 TCTGATGATGATCTCAGTCTTCAACCCAGAGGCTGCTCAGCTTCCCTCATCT 2833

RESULT 6
AX380960
LOCUS AX380960 2929 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 1 from Patent WO20212456.
ACCESSION AX380960
VERSION AX380960.1 GI:19575800
KEYWORDS

SOURCE Rattus sp.
ORGANISM Rattus sp.
Bakayota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.
TITLE Ampk-related serine/threonine kinase, designated snark
JOURNAL Patent: WO 0212456-A 1 14-FEB-2002;
1149336 ONTARIO INC. (CA)
FEATURES
source 1..2929
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"
ORIGIN
Query Match 71.2%; Score 2066.8; DB 6; Length 2929;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2488; Conservative 0; Mismatches 277; Indels 115; Gaps 14;
QY 49 GTGACCTCTGAGCCCGGGCTCAGCGCGCTGCTACTGCTGCCGACCCACCTGCACCTC 108
DB 9 GTGACCTCTGAGCTGCGGCTCTCCGCGCTGCTGCTGCTGCCGACCCCTCGCTC 68
QY 109 GCGGTCCTCCGACCATGAGTGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCG 168
DB 69 GCGGTCCCGGACCATGAGTGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCG 128
QY 169 CTTCCGCTTGGCTCGAGAGCGCCGCGCTGGCGGACGCGCTCATCAAGTGCCTA 228
DB 129 CTTCCGCTTGGCGACGAGAGCGCCGCGCTGGCGGACGCGCTCATCAAGTGCCTA 188
QY 229 AACCTCGAAGAAACAGCGCGGTGAAGCGGACCATCAACCAACCTCGGGGACC 288
DB 189 AACCTCGAAGAAACAGCGCGGTGAAGCGGACCATCAACCAACCTCGGGGACC 248
QY 289 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTTACGAGGAGTGAAGAGCAAG 348
DB 249 GCTACGAGTCTCTGAGAGCGTGGGCAAGGCACTTACGAGGAGTGAAGAGCAAG 308
QY 349 AGAGCTCGGGCGCTCTGTGGCCATCAAGTCTCAAGAAAGCAAAATCAAGATGAC 408
DB 309 AGAGCTCGGGACGCTGTGGCCATCAAGTCTCAAGAAAGCAAAATCAAGATGAC 368
QY 409 AGAGTCTGTGCAATACGAGGAGATGAGATCATGCTTCACTCAACCCCA 468
DB 369 AGAGTCTGTGCAATACGAGGAGATGAGATCATGCTTCACTCAACCCCA 428
QY 469 TCATTCGCATCATGAGTGTGTTGAGATAGCAGCAAGATGTGATTGTCAATGAGTATG 528
DB 429 TCATTCGCATCATGAGTGTGTTGAGATAGCAGCAAGATGTGATTGTCAATGAGTATG 488
QY 529 CCAAGCGAGGAGCTGTATGATTAATCAATGAGCGGCAACGCTGAGTGGAGCG 588
DB 489 CCAAGCGAGGAGCTGTATGATTAATCAATGAGCGGCAACGCTGAGTGGAGCG 548
QY 589 CCAAGGATTTCTTCCGACAGATGTGTGCTGCTGCACTACTGCAACAGAGGATG 648
DB 549 CCAAGGATTTCTTCCGACAGATGTGTGCTGCTGCACTACTGCAACAGAGGATG 608
QY 649 TTACCGAGATCTCAAGCTGGAACATCTCTTCAAGTGCATGAAAGCAATCAAGATTG 708
DB 609 TTACCGAGATCTCAAGCTGGAACATCTCTTCAAGTGCATGAAAGCAATCAAGATTG 668
QY 709 CTGACTTTGGCTCTCAACCTGTACCAAGGCAAGTCTTCAAGCTTCTGTGGGA 768
DB 669 CTGACTTTGGCTCTCTCAACCTGTATCAAAAGGCAAGTCTTCAAGCTTCTGTGGGA 728
QY 769 GCGCTCTCTAGGCTCGCTGAGATAGTCAACGAGGAGCCCTATGTGGGCCAAGGTGG 828
DB 729 GCGCTCTCTAGGCTCTCACTGAGATCTCAACGAGGAGCCCTATGTGGGCCAAGGTGG 788

QY 829 ACAAGTGTCTCTGGGCGTTCTCTGTATCATCTGTGATGAGCAATGCGACCTTTGAGC 888
DB 789 ACAAGTGTCTCTGGGCGTTCTCTGTATCATCTGTGATGAGCAATGCGACCTTTGAGC 848
QY 889 GGCAGATCATTAACCTGTGAGCAATGCAATGAGGAGCTTACGAGCGGCCA 948
DB 849 GGCAGATCATTAACCTGTGAGCAATGCAATGAGGAGCTTACGAGCGGCCA 908
QY 949 AGCGTCCGATGCTGTGGCTGATCCGCTGCTGTTAATGATGAAACCCGAGCGG 1008
DB 909 AACCTGTGATGCTGTGGCTGATCCGCTGCTGTTAATGATGAAATCCGAGCGG 968
QY 1009 CCACACTGAGATGATGACCATGATGTTGGGTCAACTGAGGTTACACACCGAGTGC 1068
DB 969 CCACACTGAGATGATGACCATGATGTTGGGTCAACTGAGGTTACACACCGAGTGC 1028
QY 1069 GGGAAACGAAAGCCCTGAGAGGTGGGACCTTATGATGACTTGGCGGGCCCA 1128
DB 1029 GGGAAACGAAAGCTTGGAGAGGTGGGACCTTATGATGACTTGGCGGGCCCA 1088
QY 1129 TGGCGACTGTTACGCTCGCTCGGCGCCCTCTGAGAAATGAGCCAGGTGCA 1188
DB 1089 TGGCGACTGTTACGCTCGCTCGGCGCCCTCTGAGAAATGAGCCAGGTGTA 1148
QY 1189 GCTTCTTCAAGCAGCAGTCCGAGAGTGAAGCACTGATCTGAGCTGAGCGGAC 1248
DB 1149 GCTTCTTCAAGCAGCAGTCCGAGAGTGAAGCAAGGACCGGGCTGAGCGGAC 1208
QY 1249 ATTCTCTTAAGAGTCCGAAAGATGATGATGCTCAAAATCTGCAAGTGAACCCG 1308
DB 1209 ATTCTCTTAAGAGTCCGAAAGATGATGATGCTCAAAATCTGCAAGTGAACCCG 1268
QY 1309 CTGAGATACCTCTTCCGCTCGGCGCAAGAGCCTTAAAGCTTCCGAAAGGCAATCTCA 1368
DB 1269 TTGAAGTACTTCTCTGCGCTTGGCAAGAACCTTCAAGCTTCCGAAAGTACTCA 1328
QY 1369 AGAAAGTCTCTACCTCTGAGGAGGTGAAGAGACCTTCAAGAACTCAAGCCG 1428
DB 1329 AGAAAGTCTCTACCTCTGAGGAGGTGAAGAGACCTTCAAGAACTCAAGCCG 1388
QY 1429 TGCTGATCTCAAGGAGAGCTGTCTCTGCTGATCTCTGCTCCAGAGAGGCAATCC 1488
DB 1389 TGCTGATCTCAAGGAGAGAGCTGTCTCTGCTGATCTCTGCTCCAGAGAGGCAATCC 1448
QY 1489 TTAAGAGTCTGAGAGGAGTGTGTTGAGTGGAGACCCCGTGAAGCAGAGTCTG 1548
DB 1449 TTAAGAGTCTGAGAGGAGTGTGTTGAGTGGAGACCCCGTGAAGCAGAGTCTG 1508
QY 1549 GGGAACTCTTGAAGCGCAGATGATGTGTTGAGTGGAGACCCCGTGAAGCAGAGTCTC 1608
DB 1509 GGGAACTCTTGAAGCGCAGATGATGTGTTGAGTGGAGACCCCGTGAAGCAGAGTCTC 1568
QY 1609 CACAGGCTTCAAGGCTCTCTCTCAACCGCAAGGCAATTTCAATGAGCAAGTTCT 1668
DB 1569 CACAGGCTTCAAGG---CGCTCAATGCAAGGCAATTTCAATGAGCAAGTTCT 1628
QY 1669 CCGGCAAGGCTTGAAGAGGAGTACCCCTGAGCACTTGGCTCCCTGAGCAACATGGCTCT 1728
DB 1626 CCGGCAAGGCTTGAAGAGGAGTACCCCTGAGCACTTGGCTCCCTGAGCAACATGGCTCT 1685
QY 1729 CTTCCCATCTGAGAGCCCGGCGCAGCCCTTCAAGGAGGCTGAGAGGAGCAGATCC 1788
DB 1686 CTTCCCATCTGAGAGCCCGGCGCAGCCCTTCAAGGAGGCTGAGAGGAGCAGATCC 1745
QY 1789 TGTCTTCGAGTCTTTGACCAATGAGCTTGGCTGAAAGTCTTCCGAAACCCCACTGA 1848
DB 1746 TGTCTTCGAGTCTTTGACCAATGAGCTTGGCTGAAAGTCTTCCGAAACCCCACTGA 1805
QY 1849 GGGGCTGTGTGCTGAGCAACCTGAGGAGGCTGAGAGGAGCTCCCTCGAGAGTCTGA 1908
DB 1806 GAGCTGTGTGTCTGAGCAACCTGAGAGGAGGCTGAGAGGAGCTCCCTCGAGAGGCTTAA 1865
QY 1909 AGGATGTGTGAGGAATCTTGTGGAGTATGCTGCTTCTTCAAGAGCTGCAAGAGG 1968

[illegible]

VERSION	AX381046.1	GI:19575868
KEYWORDS	Mus sp.	
SOURCE	Mus sp.	
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	
REFERENCE	Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.	
AUTHORS	Ampk-related serine/threonine kinase, designated snark	
TITLE	Patent: WO 0212456-A 87 14-FEB-2002;	
JOURNAL	1149336 ONTARIO INC. (CA)	
FEATURES	Location/Qualifiers	
Source	1..2027	
ORIGIN	/organism="Mus sp." /mol_type="unassigned DNA" /db_xref="taxon:10095"	
Query Match	69.7%;	Score 2022.2; DB 6; Length 2027;
Best Local Similarity	99.9%;	Pred. No. 0;
Matches 2024; Conservative	0; Mismatches	3; Indels 0; Gaps
QY	49	GTGACCTCTAGGCCGCGGGCTCAGAGCGCGCGCTGTAACGTCTCCGACCCCACTCACCTC 100
DB	1	GTGAACCTCTAGGCCGCGGGCTCAGAGCGCGCGCTGTAACGTCTCCGACCCCACTCACCTC 60
QY	109	GCGGTCCCGCACCATGAGTGTGTGGCTTACTCTCAGCGCCGAGCCAGGCTCCCTCG 168
DB	61	GCGGTCCCGCACCATGAGTGTGTGGCTTACTCTCAGCGCCGAGCCAGGCTCCCTCG 120
QY	169	CTTCGCGCTTGCGCTCTGAGAGCGCCGCGCTGCGAGCGGCTCATCAAGTGGCTTA 228
DB	121	CTTCGCGCTTGCGCTCTGAGAGCGCCGCGCTGCGAGCGGCTCATCAAGTGGCTTA 180
QY	229	AACCTCGATGAAGAAGCAGCGGCTGTGAAGCGGACCATCAAAACAAACCTTGGGACCC 288
DB	181	AACCTCGATGAAGAAGCAGCGGCTGTGAAGCGGACCATCAAAACAAACCTTGGGACCC 240
QY	289	GCTACAGATTCTGTGAGACGCTGTGGCAAGGGCACTTACGGGAAGGTGAAGAAGCAAG 348
DB	241	GCTACAGATTCTGTGAGACGCTGTGGCAAGGGCACTTACGGGAAGGTGAAGAAGCAAG 300
QY	349	AGAGCTTGGGGCGTGTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 408
DB	301	AGAGCTTGGGGCGTGTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 360
QY	409	AGAGTCTGTGTGCAATACGAGAGGAGATTGAGATCATGTCTTCACTCAACACCCCA 468
DB	361	AGAGTCTGTGTGCAATACGAGAGGAGATTGAGATCATGTCTTCACTCAACACCCCA 420
QY	469	TCAATGGCATCATGAAAGTGTGGAATAGAGCAAGATTGTGATTTGTCAATGAGATG 528
DB	421	TCAATGGCATCATGAAAGTGTGGAATAGAGCAAGATTGTGATTTGTCAATGAGATG 480
QY	529	CGAGCCGAGCGATCTGTATGATTATCATCATGAGACGGCCACCGCTGAGTAGAGCGAG 588
DB	481	CGAGCCGAGCGATCTGTATGATTATCATCATGAGACGGCCACCGCTGAGTAGAGCGAG 540
QY	589	CGAGGATTTTCTTCCGACGATCGTGTGTGGCTTGCATCTAGTCCACCAAGACGGGAT 648
DB	541	CGAGGATTTTCTTCCGACGATCGTGTGTGGCTTGCATCTAGTCCACCAAGACGGGAT 600
QY	649	TTTACCGAGATCTCAAGCTGTGAAAAATCTCTTCTAGATGCGCAATGGAACATCAAGATT 708
DB	601	TTTACCGAGATCTCAAGCTGTGAAAAATCTCTTCTAGATGCGCAATGGAACATCAAGATT 660
QY	709	CTGACTTTGGCTCTTCCAACTGTACCAAAAGGCAATTCTCCAGACGTTCTGTGGGA 768
DB	661	CTGACTTTGGCTCTTCCAACTGTACCAAAAGGCAATTCTCCAGACGTTCTGTGGGA 720
QY	769	GCCCTCTCTACGCTCTGCTGAGATGTCAACGGGAAGCCTTATGTGTGGCCAGAGGTGG 828
DB	721	GCCCTCTCTACGCTCTGCTGAGATGTCAACGGGAAGCCTTATGTGTGGCCAGAGGTGG 780

QY 529 CAGCCGAGGCGATCTGTATGATTAACATCAGTAGCGGCGACGGCTGAGTGGCGGAGC 588
 DB 481 CAGCCGAGGCGATCTGTATGATTAACATCAGTAGCGGCGACGGCTGAGTGGCGGAGC 540
 QY 589 CAGGCAATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCGCAGGACCGGATG 648
 DB 541 CAGGCAATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCGCAGGACCGGATG 600
 QY 649 TTACCGAGATCTCAGCTGGAACATCCTTTCTAGATGCCAATGGAACATCAAGATTG 708
 DB 601 TTACCGAGATCTCAGCTGGAACATCCTTTCTAGATGCCAATGGAACATCAAGATTG 660
 QY 709 CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCAAGCTTTCTGTGGA 768
 DB 661 CTGACTTTGGGCTCTCCAACTGTATCAAAAGGCAAGTTCTCTCAAGCTTTCTGTGGA 720
 QY 769 GCGCTCTCTACGCTGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCGCAGAGGTG 828
 DB 721 GCGCTCTCTACGCTGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCGCAGAGGTG 780
 QY 829 ACAGCTGCTCTTGGGGGCTTCTCTGTATCATCTGTGTGATGAGCAACATGCTTTGAGC 888
 DB 781 ACAGCTGCTCTTGGGGGCTTCTCTGTATCATCTGTGTGATGAGCAACATGCTTTGAGC 840
 QY 889 GGCAGATCATTAACAAGCTGTGAGCAAAATCAGTAACGGGCTTACCGTGAAGCGGCCA 948
 DB 841 GGCAGATCATTAACAAGCTGTGAGCAAAATCAGTAACGGGCTTACCGTGAAGCGGCCA 900
 QY 949 AGCGCTCGATGCTGTGAGCTGTATCCGGTGGCTTTAATGTGAAACCCACCGCTGCG 1008
 DB 901 AGCGCTCGATGCTGTGAGCTGTATCCGGTGGCTTTAATGTGAAATCCATCCGTGCGG 960
 QY 1009 CCACACTGAGAGATGTAGCCAGTATGTGTGCTTCACTGGGCTTACCAACCGGAGTGC 1068
 DB 961 CCACACTGAGAGATGTAGCCAGTATGTGTGCTTCACTGGGCTTACCAACCGGAGTGC 1020
 QY 1069 GGGAAACAGAAAGCCCTGCTGAGGGGTGGGACCCCTAAGTGTGACTTTGGCGGCGCTCA 1128
 DB 1021 GGGAAACAGAAAGCTCTGCGAAGGGGTGGGACCCCTAAGTGTGACTTTGGCGGCGCTCA 1080
 QY 1129 TGGGGAATGCTGCTGCTGCTGCTGCGGCCCTCTCTGAGATGAGCCAAAGTGTGCA 1188
 DB 1081 TGGGGAATGCTGCTGCTGCTGCTGCGGCCCTCTCTGAGATGAGCCAAAGTGTGCA 1140
 QY 1189 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGGAACATCTGTCTGTGGCTGTGAGCGGCAAC 1248
 DB 1141 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGGAACATCTGTCTGTGGCTGTGAGCGGCAAC 1200
 QY 1249 ATTCTCTTAAGAAATCCCGAAAGAGATGATGAGCTCAAAATCTGCAAGTGTGACCGCG 1308
 DB 1201 ATTCTCTTAAGAAATCCCGCAAGAGATGATGAGCTCAAAATCTGCAAGTGTGACCGAG 1260
 QY 1309 CTGAGATATCTTCTTCTGCGCCCTGCAAGAGCAGCTTAAAGCTTCCGAAAGCATTTCTCA 1368
 DB 1261 TTGAAAGATATCTTCTTCTGCGCCCTGCAAGAGCAGCTTAAAGCTTCCGAAAGCATTTCTCA 1320
 QY 1369 AGAAAAAGTCTCTTACTCTGTCAGGGAGGTATCAGAGAGACCTCTGAGAACTCAGACCGG 1428
 DB 1321 AGAAAAAGTCTCTTACTCTGTCAGGGAGGTATCAGAGAGACCTCTGAGAACTCAGACAG 1380
 QY 1429 TGCCTGATATCTCAGGGGAGCTGCTGCTGTATCCCTGCTCCAAAGAAAGGCAATCC 1488
 DB 1381 TGCCTGATATCTCAGGGGAGCTGCTGCTGTATCCCTGCTCCAAAGAAAGGCAATTC 1440
 QY 1489 TTAAAGATCTCAGCAGCGTGAATCTGTACTCTCTCAGAGCGCCAGCAGTCTG 1548
 DB 1441 TTAAAGATCTCAGCAGCGTGAATCTGTACTCTCTCAGAGCGCCAGCAGTCTG 1500
 QY 1549 GGGAACTCTTAAGACGCAAGTGTGTTGTGAGTGGGAGCCCGCTGAGCAGAAAGTCTC 1608
 DB 1501 GGGAACTCTTAAGACGCAAGTGTGTTGTGAGTGGGAGCCCGCTGAGCAGAAAGTCTC 1560
 QY 1609 CACAGGCTTACGAGGCTCTCTCTCCACCGCAAGGCAATTTCAAACTCAATGCGCAAGTCT 1668

DB 1561 CACAGCTTACAGG---CGCTTCATGCGAAGGCAATCTCAAACTCAATGCGCAAGTTT 1617
 QY 1669 CCGGACAGGCTTTAAGAGGACATACCCCTTAGCACTTTGCTCTCTGAGCAATGCGCT 1728
 DB 1618 CCGGACAGGCTTTAAGAGGACATACCCCTTAGCACTTTGCTCTCTGAGCAATGCGCT 1677
 QY 1729 CTTCCCATCTCTGAGACCCCGGCGCAGCGCCCTCAGGGGCTGTGAGTGAAGACAGATCC 1788
 DB 1678 CCGCTCATCTTACAGCCCGGCGCAGCGCTCTCTGAGGCTGTGAGTGAAGACAGCATCC 1737
 QY 1789 TGTCTCCGAGTCTTTGACCAATTTGACTGTGCTGTAACGCTCTTCCGAAACCCCACTGA 1848
 DB 1738 TGTCTCCGAGTCTTTGACCAATTTGACTGTGCTGTAACGCTCTTCCGAAACCCCACTGA 1797
 QY 1849 GGGGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTGAAGCGCTCTCTGAGAAAGTCTGA 1908
 DB 1798 GGAAGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTGAAGCGCTCTCTGAGAAAGTCTGA 1857
 QY 1909 AGGATGTGTGTGTGTGTGAGCAACCTGAGGGGATGTGCTTTTCTGTGACAGATCTGCAAGG 1968
 DB 1858 AGGATGTGTGTGTGTGTGAGCAACCTGAGGGGATGTGCTTTTCTGTGACAGATCTGCAAGG 1917
 QY 1969 TGACTGAGGCTTACAGCAAGCCCTAGAGATCTGTCTCAAGCTCAGCTGAGGAGGAGA 2028
 DB 1918 TGACAGAGGCTTACAGCAAGCCCTAGAGATCTGTCTGAGAGCTCAGCTGAGGAGGAGG 1977
 QY 2029 TGGTCCCTAGT-ATGGGATAGGCTCTGAGA-GGGTTTGAGAGAAAC 2075
 DB 1978 CAGTGCCCGAGTATGGGGTGAAGCTTTAGAGGGGTTTGAGAGAAAC 2026

RESULT 9
 AC125887
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
 *** 13 unordered pieces.
 AC125887
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 1 (bases 1 to 291762)
 Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S,, Adams,C,, Alder,J.,
 Allen,C., Allen,H., Alibrooks,S., Amin,A., Anguiano,D.,
 Araya-Lebeche,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Bissale,K., Blair,J., Blankenburg,K., Blythe,P., Brown,M.,
 Bryant,N., Buay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cre,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,B., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Bayes,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Frazer,C.M., Gabiela,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gaili,C., Grady,M., Guerra,W., Guvava,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
 Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kover,C.,
 Kowitz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,D., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensbuewa,L., Louieged,M., Lozano,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Mallory,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mitoavljivic, A., Miner, G., Minya, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachukwu, O., Okunolu, G., Olarnpungsoon, A., Pal, S., Paris, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldinger, A., Popovic, D., Primus, B., Pu, L., L., Piazzi, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Sytek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Velas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
Direct Submission
2 (bases 1 to 291762)

JOURNAL
Unpublished

REFERENCE
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

AUTHORS
3 (bases 1 to 291762)

JOURNAL
Rat Genome Sequencing Consortium.
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Nov 15, 2002 this sequence version replaced gi:23269681. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GMLD
Center Clone name: CH230-274J3
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 230853 bases at least Q30
Consensus quality: 233437 bases at least Q20
Estimated insert size: 233445; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 70277: contig of 70277 bp in length
* 70278 70377: gap of unknown length
* 70378 70378: contig of 24654 bp in length
* 95032 95131: gap of unknown length
* 95132 108956: contig of 13825 bp in length
* 108957 109056: gap of unknown length
* 109057 271697: contig of 162641 bp in length
* 271698 271797: gap of unknown length
* 271798 273026: contig of 1229 bp in length
* 273027 273126: gap of unknown length
* 273127 274531: contig of 1405 bp in length
* 274532 274631: gap of unknown length
* 274632 276287: contig of 1656 bp in length
* 276288 276387: gap of unknown length
* 276388 277904: contig of 1517 bp in length
* 277905 278004: gap of unknown length
* 278005 279523: contig of 1519 bp in length
* 279524 279623: gap of unknown length
* 279624 280724: contig of 1101 bp in length
* 280725 280824: gap of unknown length
* 280825 284658: contig of 3834 bp in length
* 284659 284758: gap of unknown length
* 284759 287234: contig of 2476 bp in length
* 287235 291762: contig of 4428 bp in length.
* 287335 291762: contig of 4428 bp in length.

FEATURES
source
1. 291762
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-274J3"
70378..75596
misc_feature
/note="wgs contig"
95132..96554
misc_feature
/note="wgs contig"
109057..110574
misc_feature
/note="wgs contig"

ORIGIN
Query Match 54.9%; Score 1593.2; DB 2; Length 291762;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

55 TCTGAGCCCGGCGTCAAGCGCGGCTGCTACTGCTGCGGCGGCGGCGGCGGCTC 114
274674 TCTGAGCCCGGCGTCTCCGCGGCTGCTACTGCTGCGGCGGCGGCGGCTC 274733

115 CCGGACACATGAGTGGTGGCTTACTCCAGGCGCGGAGCGAGCTCCGCGCTCCG 174
274734 CCGGACACATGAGTGGTGGCTTACTCCAGGCGCGGAGCGAGCTCCGCGCTCCG 274793

175 CCTGCGCTGAGAGCGCCCGCGCTGCGGAGCGGCTCATCAAGTGGCTTAACTC 234
274794 CCTGCGCTGAGAGCGCCCGCGCTGCGGAGCGGCTCATCAAGTGGCTTAACTC 274853

235 TGATGAAGAGCGGCGCTGAAGCGGCGCATCAACAACCACTCGCGGCGCTACG 294
274854 TGATGAAGAGCGGCGCTGAAGCGGCGCATCAACAACCACTCGCGGCGCTACG 274913

295 AGTTCTGAGAGCGCTGGGCAAGGCGACCTTAAGGAGGAGTGAAGAGGAGAGGCT 354
274914 AGTTCTGAGAGCGCTGGGCAAGGCGACCTTAAGGAGGAGTGAAGAGGAGGCT 274973

355 CGGGCGCTGTGGCGCATCAAGTCAATCAGAAAGCAAAATCAAAATGAGCAGATC 414
274974 CGGGCGCTGTGGCGCATCAAGTCAATCAGAAAGCAAAATCAAAATGAGCAGATC 275033

415 TGCTGCATACGAGGAGGAGATGAGTCAATGCTTCACTCAACACCCCACTATTG 474
275034 TGCTGCATACGAGGAGGAGATGAGTCAATGCTTCACTCAACACCCCACTATTG 275093

Db 644 TGCAATATGGCATCAGAAACAGAGTTGTCCACCGAGATCTCAAGCTGGAGAAATCCTCT 703
 Qy 682 TAGATGCAATGAAACATCAAGATTTGCTGACTTTGGCTCTCCACCTGTACCAAG 741
 Db 704 TGAATGCAATGGAATATCAAGATTTGCTGACTTTGGCTCTCCACCTGTACCAAG 763
 Qy 742 GCAAGTTCTTCAGACGTTCTGTGGAGCCCTCTCAAGCTCTGCTGAGATGATCAAG 801
 Db 764 GCAAGTTCTTCAGACATTTCTGTGGAGCCCTCTCAAGCTCTGCTGAGATGATCAAG 823
 Qy 802 GGAAGCCCTATATGTTGGGCTCCAGAGGTGACAGCTGTCTGTGGGCTTCTCTGTACATCC 861
 Db 824 GGAAGCCCTATACAGAGGCTCCAGAGGTGACAGCTGTCTGTGGGCTTCTCTGTACATCC 883
 Qy 862 TGGTGCATGAGCACCATGCTTTTGAACGGGAGAGATCAATAAACACTGGTGAAGAAATCA 921
 Db 884 TGGTGCATGAGCACCATGCTTTTGAATGGGAGATGACATTAAGATCTTATGTAACAGATCA 943
 Qy 922 GTAAACGGGGCTTACCGTGAAGCCGCCCAAGCCGTCCAGTCTGTGGCTGATCCGGTGGC 981
 Db 944 GCAACGGGGCTTACCGGGAGCCACCTAAACCTCTGATGCTGTGGCTGATCCGGTGGC 1003
 Qy 982 TGTATATGTTGAACCCCAACCTCTGTGGGCTCACTGAGAGATGTAGCCATTTGGTGG 1041
 Db 1004 TGTATATGTTGAACCCCAACCTCTGTGGGCTCACTGAGAGATGTGGCCATGCTGTGGG 1063
 Qy 1042 TCAATCGGGGTTACACCAACCGGAGTCCGGGGAACAGAAAGCCCTGCTGAGGGTGGGACCC 1101
 Db 1064 TCAATCGGGGTTACGCGCACCGAGTGGGAGAGAGAGGCTCCGATGAGGGTGGGACCC 1123
 Qy 1102 CTAGTGTGACTTTGGCCGGGCTCTCAATGCGGAGCTGTGTAAGTGTGCTCTCCGCGCCCC 1161
 Db 1124 CTGCAATGACTCTGCGCGGCTCTCAATGCTGACTGTGCTCGGCGTTCTCTCCGCGCCCC 1183
 Qy 1162 TCTCTGAGATGAGCAAGGCTGTGACGCTTCTTCAAGCAGACGCTGCGGAGGTGAA 1221
 Db 1184 TCTCTGAGATGAGGAGGCTGTGACGCTTCTTCAAGCAGACGCTGCTGGGAGGAA 1243
 Qy 1222 GCACTGATACCTGGGCTGAGAGGCAATTTCTTAAAGATCCCGAAGAGAGATGACA 1281
 Db 1244 GCACTGATACCTGGGCTGAGAGGCAATTTCTTAAAGATCCCGAAGAGAGATGACA 1303
 Qy 1282 TGGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTTCTCGCCCTGGCAAGACA 1341
 Db 1304 TGGCTCAAGTCTCTCAAGGTGACCCGGCTGAGATACCTTCTCGCCCTGGCAAGACA 1363
 Qy 1342 GCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTTCAAGCTGTCAGGGAGGTAC 1401
 Db 1364 ACCTCAAGGTGCAAGAGGCAATTTCTCAAGAAAGTGTCAAGCTTCTGCAAGAGGGTAC 1423
 Qy 1402 AGAGAGACCTGCAAGAACTGACACCGGTGCTGATATCTCAAGGGGAGGCTTCTCTGCTG 1461
 Db 1424 AGAGAGACCTGCAAGAGTCAAGGCTCAAGCTTCTGCAAGGGGAGGCTCTCTCTCTCTCT 1477
 Qy 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTTAAAGATCTGACAGGCTGAATCTGCTTACT 1521
 Db 1478 ---CCCTGCTCCCAAGGAAAGGCAATCTTCAAGAAAGGCTGACAGGCTGCTGCTACT 1534
 Qy 1522 ACTCTCTCCAGAGCCGAGAGTCTGAGGAACTCTTAAAGCCGCAAGTATGTTTGTGA 1581
 Db 1535 ACTCTCTCCGAGCCGAGAGTCTGAGGAACTCTTAAAGCCGCAAGTATGTTTGTGA 1594
 Qy 1582 GTGGGAGCCCGTGAAGAGATGCTCAAGGCTTCAAGGCTTCTCTCAAGGCAAGG 1641
 Db 1595 GTGGGAGTCCCAAGAGAGAGGCTCGCAAGGCTTCAAGGCTTCTCTCAAGGCAAGG 1654
 Qy 1642 GCATTCTCAAACTCAATGCAAGTCTCTCCGCAAGGCTTGAAGAGCACTACCCCTTACA 1701
 Db 1655 GCATTCTCAAACTCAATGCAAGTCTCTCCGCAAGGCTTGAAGAGTCTGCGGCCCCCACTA 1714
 Qy 1702 CTTTGTGCTCCCTGAGCAACTGTGCTCTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1761
 Db 1715 CTTTGTGCTCCCTGAGCAACTGTGCTCTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1774

Qy 1762 CAGGGGCTGTAGTGAAGACAGATCTGCTCTCCAGAGTCTTTGACCAATTTGACTTGC 1821
 Db 1775 CAGGGGCTGTAGTGAAGACAGATCTGCTCTCCAGAGTCTTTGACCAATTTGACTTGC 1834
 Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGTGAACAACCTGAGGGG 1881
 Db 1835 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTGTGTGTGAACAACCTGAGGGG 1894
 Qy 1882 TTGAGCAGCCCTCCGAGAG-----GTGAGAGCGATGAGTGGCAGGAATCT 1929
 Db 1895 TTGAGAGCCCTCCGAGAGGCTCCGAGAGCTGAGAGCTTGGCGGAGATCTT 1954
 Qy 1930 TGGGGGATAGCTGCTTCTTCTGACAGACTGCCAAGAGTGAATCTGACAGCAAG 1989
 Db 1955 TGGGGGATAGCTGCTTCTTCTGACAGACTGCCAAGAGTGAATCTGACAGCAAG 2014
 Qy 1990 CCTGAAGATCTGCTCAAGAGTCACTGAGAGAGGAGATGTCCTTATGATGGGTAG 2049
 Db 2015 CACTGAGAGTCTGCTCAAGAGTCACTGAGAGAGGAGATGTCCTTATGATGGGTAG 2073
 Qy 2050 GCTCTGAGAGGTTTGAAGAGAACCTGGGTGGATTTCTCC 2092
 Db 2074 GCTCTGAGAGTCACTGCTGTCACCCGAGGAGATGCTTC 2116

RESULT 11
 LOCUS CQ782778 3395 bp DNA linear PAT 17-MAR-2004
 DEFINITION Sequence 2918 from Patent EP1396543.
 ACCESSION CQ782778
 VERSION CQ782778.1 GI:45502721
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Oca, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
 Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
 Koga, H.
 TITLE Primers for synthesizing full length cDNA clones and their use
 JOURNAL Patent: EP 1396543-A 2918 10-MAR-2004;
 Research Association for Biotechnology (JP)
 FEATURES
 source
 1..3395
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 118..2004
 /note="unamed protein product"
 CDS
 /codon_start=1
 /protein_id="CAF6008.1"
 /db_xref="GI:45502722"
 /translation="MESLVEFARBSGPTSAEILARPLAEGILKSPKPLMKKQAVRHH
 HKNRHPRVERELTIGKTYGKYGKAGSSGILVATISIRDKTKDODIMHRRBIE
 IMSLNHPHIIAHEVPSNSKIYIVMEYSRGLUYIYSRQLSERBAHFPRQIV
 SAVHYCHQNVHRDLKLENLILDPANIKIADGLSNLHQGFLOTPCGSPLYASP
 EIVNGKPYTPEDVSWISGLVLYLVHGTGTPFDGDKHLVQKSNQAYRBPSPSDA
 CGLIRMLINPTFRATLEDAVASHMWNMGVATVGOEAPHEGSHGSDSARASMA
 MLRRSRPLLENGAKVCSFPKQAPGGGSTTGLERHSLKSKSKENDMAOSLSDTA
 DDPARHPKSNILKIPKGIILKKYASAEVGOEDPELSPIPASGGOAAILPKGIILK
 KPRORESGYISSPERSGELLDGDVFGSDPREGQRPQASGILHRRKGLIKANGK
 SQTLELAAPTFRESLDELAPRPLAASRPSGAVSDSILSSFPDLIDLPRLPRLP
 PLRGCVCVDNI/TLBEPSPSPGSLRMRQDPLGDSFSLDQEVATAYRQALRVC
 SKLT"
 ORIGIN
 Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
 Best Local Similarity 81.7%; Pred. No. 0;
 Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
 Qy 82 CTACTGTGCCGACCACTCACTGCGGCTCCCGCAGCATGAGTGGTGGCTTAC 141

Db 77 |T|A|T|G|A|T|T|C|C|C|C|C|C|C|T|T|G|T|C|A|C|T|C|T|G|C|T|G|C|A|T|G|A|G|T|C|G|T|T|T|C|G| 136
Qy 142 |T|C|A|G|C|G|C|G|G|A|G|C|C|A|G|G|C|T|C|C|T|G|G|C|T|T|G|C|C|C|T|G|G|C|T|G|A|G|C|C|C|G|G|C|G|C| 201
Db 137 |C|G|C|G|C|C|C|T|C|G|G|C|C|C|C|C|A|C|T|C|C|T|C|G|G|C|G|C|A|G|A|G|T|A|-----|G|C|C|G|G|C|G|C| 184
Qy 202 |T|G|G|G|G|A|G|C|G|G|C|T|C|A|T|G|T|G|G|C|T|T|A|A|C|T|G|A|T|G|A|A|G|A|G|G|G|T|G|A|G|C|G|C| 261
Db 185 |T|G|G|G|G|A|A|G|G|G|C|T|A|T|C|A|G|T|G|G|C|C|A|G|C|C|C|T|A|T|A|T|A|A|G|A|G|G|C|G|T|G|A|G|C|G|G|C| 244
Qy 262 |A|C|C|A|T|C|A|C|A|A|C|A|C|C|T|G|G|G|C|A|C|G|C|T|A|G|A|T|T|C|T|G|G|A|G|C|G|T|G|G|C|A|G|G|C|A| 321
Db 245 |A|C|C|A|C|C|A|G|A|G|C|A|C|C|T|G|G|G|C|A|C|C|G|C|T|A|G|A|T|T|C|T|G|G|A|G|C|C|T|G|G|C|A|G|G|C|A| 304
Qy 322 |C|T|A|G|G|G|A|G|T|G|A|G|A|G|G|C|A|G|A|G|A|G|C|T|G|G|G|G|G|G|T|G|G|T|G|G|C|C|A|A|G|T|C|A| 381
Db 305 |C|T|A|G|G|G|A|G|T|G|A|G|A|G|G|C|G|G|G|G|A|G|C|T|G|G|G|G|G|G|C|G|T|G|G|G|C|C|A|A|G|T|C|A| 364
Qy 382 |T|C|A|G|G|A|A|G|A|C|A|A|A|T|C|A|A|A|A|T|G|A|G|C|A|G|A|T|T|G|T|G|C|C|A|T|A|T|A|C|G|A|G|G|A|G|T|T|G|A| 441
Db 365 |T|C|C|G|A|A|G|G|A|C|A|A|A|A|T|C|A|A|A|A|T|G|A|G|C|A|A|G|A|T|T|G|A|T|G|C|A|T|A|C|G|A|G|G|A|G|A|T|T|G|A| 424
Qy 442 |T|C|A|T|G|T|C|T|C|A|C|A|C|C|C|C|C|A|C|A|T|C|A|T|T|G|C|C|A|T|C|A|T|G|A|G|T|T|T|G|A|A|T|G|A| 501
Db 425 |T|C|A|T|G|T|C|A|C|T|C|A|C|C|C|C|C|A|C|A|T|C|A|T|T|G|C|C|A|T|C|A|T|G|A|A|G|T|T|T|G|A|A|A|C|A|G|A| 484
Qy 502 |G|C|A|A|G|T|T|G|A|T|T|G|C|A|T|G|A|T|G|C|A|G|C|C|G|A|G|G|C|A|T|C|T|G|A|T|G|A|T|T|A|C|A|T|C|A|G|T|G| 561
Db 485 |G|C|A|A|G|T|G|T|G|A|T|G|C|A|T|G|A|T|G|C|A|G|C|C|G|G|G|G|A|C|T|T|T|A|T|G|A|C|A|T|C|A|C|A|G|C| 544
Qy 562 |A|G|G|G|C|C|A|C|G|G|C|T|G|A|G|A|G|G|G|G|A|C|C|C|A|G|G|C|A|T|T|T|T|C|C|G|A|C|A|A|T|C|G|T|G|T|C|C|C| 621
Db 545 |A|G|C|G|C|A|G|C|A|G|T|G|A|G|G|G|C|C|A|G|A|G|T|T|T|T|C|C|G|G|A|G|A|T|C|C|T|C|T|C|G|C|G| 604
Qy 622 |T|G|C|A|T|C|T|G|C|C|A|C|A|G|A|C|G|G|A|T|C|G|T|T|C|A|C|C|G|A|G|A|T|C|T|A|G|C|T|G|G|A|A|A|A|C|A|C|T|C|T| 681
Db 605 |T|G|C|A|T|T|T|G|C|C|A|C|A|G|A|G|A|T|T|G|T|C|C|A|C|C|G|A|G|A|T|C|T|A|G|C|T|G|G|A|G|A|A|C|A|T|C|T| 664
Qy 682 |T|A|G|T|G|C|A|T|G|A|A|A|C|A|A|G|A|T|G|C|T|G|A|C|T|T|T|G|G|C|T|T|C|C|A|C|C|T|G|T|A|C|A|C|A|A|G| 741
Db 665 |T|G|A|T|G|C|C|A|A|T|G|G|A|A|T|A|T|C|A|A|G|A|T|G|T|G|A|C|T|T|C|G|G|C|T|C|C|A|C|C|T|T|A|C|A|C|A|T|G| 724
Qy 742 |G|C|A|G|T|C|T|C|C|A|C|A|G|T|T|C|T|G|G|A|G|C|C|T|C|T|A|C|G|C|T|G|C|T|G|A|T|A|G|T|C|A|G| 801
Db 725 |G|C|A|G|T|C|T|G|C|A|C|A|T|T|C|T|G|G|A|G|C|C|C|C|C|T|A|T|G|C|T|C|C|C|A|G|A|G|A|T|T|G|C|A|T|G| 784
Qy 802 |G|A|A|G|C|C|T|A|T|G|T|G|G|C|C|A|G|A|G|T|G|A|C|A|G|C|T|G|T|C|T|G|G|G|G|T|C|C|T|G|T|A|C|A|T|C| 861
Db 785 |G|A|A|G|C|C|T|A|C|A|G|G|C|C|A|G|G|T|G|A|C|A|G|C|T|G|T|C|C|T|G|G|T|G|T|C|C|T|T|A|C|A|T|C| 844
Qy 862 |T|G|T|G|C|A|T|G|C|A|C|C|C|T|T|T|G|A|C|G|G|C|A|G|A|T|C|A|T|A|A|A|C|A|C|T|G|T|G|A|G|C|A|A|T|C|A| 921
Db 845 |T|G|G|G|C|A|T|G|C|A|C|C|C|C|T|T|T|G|A|T|G|G|G|C|A|T|G|C|A|T|A|A|G|A|T|C|T|A|G|T|G|A|A|A|C|A|G|A|T|C|A| 904
Qy 922 |G|T|A|C|G|G|G|G|C|T|A|C|C|G|G|A|G|C|C|C|C|A|A|G|C|C|T|C|C|A|T|G|C|T|G|G|G|C|T|G|A|T|C|C|G|G|G|C| 981
Db 905 |G|C|A|C|G|G|G|G|C|T|A|C|C|G|G|A|G|C|C|T|A|A|C|C|C|T|G|A|T|G|C|T|G|G|G|C|T|G|A|T|C|C|G|G|G|C| 964
Qy 982 |T|G|T|A|T|G|T|G|A|C|C|C|C|C|G|G|G|C|A|C|A|C|T|G|A|G|A|T|G|A|G|C|A|G|A|T|G|T|G|G|G| 1041
Db 965 |T|G|T|A|T|G|T|G|A|C|C|C|C|C|G|G|G|C|A|C|C|T|G|A|G|A|T|G|A|G|C|C|A|G|A|T|G|T|G|G|G| 1024
Qy 1042 |T|C|A|A|C|T|G|G|G|G|T|A|C|A|C|C|G|G|A|T|C|G|G|G|A|A|C|G|A|A|G|C|C|T|G|T|G|A|G|G|T|G|G|C|A|C| 1101
Db 1025 |T|C|A|A|C|T|G|G|G|G|T|A|C|C|C|C|C|G|G|A|T|G|G|G|A|G|A|G|G|G|C|C|C|C|C|A|T|A|G|G|G|T|G|G|C|A|C| 1084
Qy 1102 |C|T|A|G|T|G|T|G|A|C|T|T|G|C|C|G|G|G|C|C|T|C|A|T|G|C|G|G|A|C|T|G|G|T|A|G|T|G|T|A|G|T|C|C|T|C|G|C|C|C|C|C| 1161
Db 1085 |C|T|G|C|A|G|T|G|A|C|T|G|C|C|C|G|G|C|C|C|C|C|A|T|G|G|T|G|A|C|T|G|G|C|C|G|G|G|T|T|C|C|C|G|C|C|C|C|C| 1144
Qy 1162 |T|C|C|G|G|A|A|T|G|A|G|C|A|G|G|T|G|G|C|T|T|C|A|G|C|A|G|C|A|G|T|C|C|G|G|A|G|T|G|A| 1221

Db 1145 |T|C|C|T|G|A|A|T|G|G|G|G|C|C|A|G|G|T|G|C|A|G|C|T|T|C|T|C|A|G|C|A|G|C|A|T|G|C|A|C|T|G|T|G|G|G|A| 1204
Qy 1222 |G|C|A|C|T|G|A|C|T|G|G|G|C|T|G|A|G|G|G|C|A|C|A|T|T|C|T|T|A|A|A|G|T|C|C|G|A|A|G|A|A|T|G|A| 1281
Db 1205 |G|C|A|C|A|C|C|C|T|G|G|C|T|G|A|G|G|C|C|A|G|C|A|T|G|C|T|C|A|A|A|G|T|C|C|G|C|A|A|G|A|G|A|T|G|A| 1264
Qy 1282 |T|G|G|C|T|A|A|A|T|C|T|G|C|A|G|T|G|A|C|C|G|G|C|T|A|G|A|T|A|C|C|T|T|T|C|G|G|C|C|T|G|G|C|A|A|G|C|A| 1341
Db 1265 |T|G|G|C|C|A|G|T|C|T|C|A|C|A|G|T|G|A|C|G|G|C|T|A|T|G|A|C|A|T|G|C|C|C|A|T|G|C|C|T|G|G|C|A|A|G|C|A| 1324
Qy 1342 |G|C|T|T|A|G|C|T|C|C|A|A|A|G|C|A|T|T|C|A|G|A|A|A|A|A|G|T|C|T|C|T|A|C|C|T|C|G|T|C|A|G|G|G|A|G|T|A|C| 1401
Db 1325 |A|C|C|T|C|A|G|C|T|C|C|A|A|A|G|G|C|A|T|T|C|A|A|A|A|A|A|G|T|G|T|C|A|G|C|T|C|G|C|A|A|G|G|G|T|A|C| 1384
Qy 1402 |A|G|A|G|A|C|C|T|C|A|G|A|A|C|T|A|C|A|C|G|G|T|G|C|T|G|A|T|C|T|C|A|G|G|C|A|G|C|T|G|T|C|C|T|G|C|T|G| 1461
Db 1385 |A|G|A|G|A|C|C|C|T|C|G|A|A|C|T|C|A|G|C|C|A|T|C|C|T|G|G|A|C|C|C|A|G|G|C|A|G|G|C|T|G|C|C|-----| 1438
Qy 1462 |T|A|T|C|C|T|G|C|T|C|C|A|G|A|A|A|G|C|A|T|C|T|T|A|A|A|G|T|T|C|G|A|C|A|G|C|T|G|A|A|T|C|T|G|T|T|A|C|T| 1521
Db 1439 |---|C|C|T|G|C|T|C|C|A|A|A|G|G|C|A|T|T|C|T|C|A|A|A|G|C|C|C|C|A|C|A|C|G|G|A|G|T|G|G|T|A|C|T| 1495
Qy 1522 |A|C|T|C|C|T|C|C|A|G|A|C|C|C|A|G|G|A|G|T|C|T|G|G|G|A|A|C|T|T|T|A|A|C|G|C|A|G|C|A|T|G|T|T|T|G|A| 1581
Db 1496 |A|C|T|C|C|T|C|C|C|A|G|A|C|C|C|A|G|T|G|A|T|C|T|G|G|A|G|C|T|T|T|G|A|G|C|A|G|C|A|G|C|G|T|T|T|G|A| 1555
Qy 1582 |G|T|G|G|G|A|C|C|C|G|T|G|A|G|C|A|A|G|T|T|C|C|A|C|A|G|G|T|T|C|A|G|G|C|T|C|T|C|C|A|C|C|G|C|A|G| 1641
Db 1556 |G|T|G|G|G|A|T|C|C|C|A|A|G|A|C|A|A|A|G|C|C|C|C|G|C|A|G|G|T|T|C|A|G|G|C|T|C|C|T|C|A|T|G|C|A|A|G| 1615
Qy 1642 |G|C|A|T|T|C|C|A|A|C|T|C|A|A|T|G|G|A|G|T|T|C|C|C|G|C|A|C|A|G|C|T|T|A|G|A|A|G|C|A|T|A|C|C|C|T|A|G|C|A| 1701
Db 1616 |G|C|A|T|C|T|C|A|A|C|T|C|A|A|T|G|G|A|A|G|T|T|C|C|C|A|C|A|C|A|G|C|T|T|G|A|G|C|T|G|C|G|G|C|C|C|C|A|C|A| 1675
Qy 1702 |C|C|T|T|G|G|C|C|C|T|G|A|C|C|A|A|C|T|G|G|C|C|C|T|C|C|C|A|C|C|T|G|A|G|C|C|G|G|C|C|C|A|G|C|C|C|C|T| 1761
Db 1676 |C|C|T|T|G|G|C|C|C|C|T|G|A|G|A|A|C|T|G|G|C|C|C|A|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C| 1735
Qy 1762 |C|A|G|G|G|C|T|G|T|G|A|G|A|G|A|C|A|G|A|T|C|T|G|T|C|C|C|G|A|G|T|C|C|T|T|G|A|C|A|A|T|T|G|A|C|A|T|T|G|C| 1821
Db 1736 |C|A|G|G|G|C|T|G|T|G|A|G|A|G|A|C|A|G|A|T|C|T|G|T|C|T|G|A|G|T|C|T|T|G|A|C|A|G|T|G|A|C|A|T|T|G|C| 1795
Qy 1822 |C|T|G|A|A|C|T|T|T|C|C|G|A|A|C|C|C|C|A|T|G|A|G|G|G|C|T|G|T|G|T|C|T|G|T|G|A|C|A|A|C|T|G|A|G|G|G|C| 1881
Db 1796 |C|T|G|A|A|C|G|C|T|C|C|A|G|A|C|C|C|C|C|A|C|T|G|G|G|G|C|T|G|T|G|T|C|T|G|T|G|A|C|A|A|C|T|C|A|C|G|G|G|C| 1855
Qy 1882 |T|T|G|A|G|C|C|C|C|C|C|C|C|A|G|A|G|-----|G|T|C|A|A|G|C|A|T|G|T|G|G|C|A|G|A|A|T|C|T| 1929
Db 1856 |T|T|G|A|G|A|G|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C|C| 1915
Qy 1930 |T|G|G|G|G|A|T|G|A|C|T|G|T|T|C|T|G|A|C|A|G|C|T|G|C|C|A|G|A|G|T|G|A|C|T|G|C|A|G|C|T|T|A|C|A|G|A|G| 1989
Db 1916 |T|G|G|G|G|A|C|A|G|C|T|G|T|T|C|C|C|T|G|A|C|A|G|C|T|G|C|C|A|G|A|G|G|T|G|A|C|A|G|C|A|C|C|T|A|C|C|G|A|C|A|G| 1975
Qy 1990 |C|C|T|A|G|A|A|T|C|T|G|C|C|A|A|A|C|T|C|A|G|C|T|G|A|G|A|A|G|G|A|G|T|G|C|C|T|A|T|A|T|G|G|G|T|A|G| 2049
Db 1976 |C|A|C|T|G|A|G|G|T|G|T|C|T|C|A|A|A|C|T|C|A|C|C|G|A|T|G|A|G|T|G|G|C|A|T|T|G|C|C|C|C|A|G|-----|C|C|G|G|T|C|A|G| 2034
Qy 2050 |G|C|T|C|T|G|A|G|G|G|T|T|G|C|A|G|A|G|A|C|C|C|T|G|G|T|G|G|A|T|T|C|T|C|C| 2092
Db 2035 |G|C|T|C|C|A|G|A|T|G|C|A|G|C|T|G|T|T|G|C|A|C|C|C|C|G|A|G|G|A|G|A|T|G|C|T|T|C| 2077

RESULT 12
BD127309
LOCUS BD127309 3395 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127309
VERSION BD127309.1 GI:23222254
KEYWORDS JP 2002017375-A/2740.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 3395)
Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE
JOURNAL
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2740 22-JAN-2002;
OS Homo sapiens (human)
PN JP 2002017375-A/2740
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (118)..(2001).
1.3395
Location/Qualifiers
1.3395
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
QY 82 CTAAGTCTGCTGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 141
DB 77 CTAAGTCTGCTGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 136
QY 142 TCCAGGCGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 201
DB 137 CGCGGCGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 184
QY 202 TGGCGGACCGGCTCATGAGTGGCTTAACCTGATGAAGACGAGCGGTGAAGCGGC 261
DB 185 TGGCGGACCGGCTCATGAGTGGCTTAACCTGATGAAGACGAGCGGTGAAGCGGC 244
QY 262 ACCATCAACAACAACCTGCGGACCGCTACGAGTTCTTGAGACGCTGCGGACGCA 321
DB 245 ACCACCAACAACAACCTGCGGACCGCTACGAGTTCTTGAGACGCTGCGGACGCA 304
QY 322 CTTACGGGAAGGTGAAGACGAGAGAGCTGGGGCGTCTGTGGCCATCAAGTCCA 381
DB 305 CTTACGGGAAGGTGAAGACGAGAGAGCTGGGGCGTCTGTGGCCATCAAGTCCA 364
QY 382 TCAGGAAGAACAACAATGAGAGAGAGTGTGTCATCATACGAGGAGGAGATTGGA 441
DB 365 TCAGGAAGAACAACAATGAGAGAGAGTGTGTCATCATACGAGGAGGAGATTGGA 424
QY 442 TCATGTCTTCACTCAACCAACCCCAATCATTCATCAGAAAGTTTGAAGATGCA 501
DB 425 TCATGTCTTCACTCAACCAACCCCAATCATTCATCAGAAAGTTTGAAGATGCA 484
QY 502 GCAAGATTGTGATGTATGAGAGATGCGACCGGAGGATCTGTATGATTAATCATAGT 561
DB 485 GCAAGATTGTGATGTATGAGAGATGCGACCGGAGGATCTGTATGATTAATCATAGT 544
QY 562 AGCGGCAACGCGCTGAGTGAAGCGGACGCAATTTCTTCCGACAGATCGTCTGCC 621
DB 545 AGCGGCAACGCGCTGAGTGAAGCGGACGCAATTTCTTCCGACAGATCGTCTGCC 604
QY 622 TGCACTAAGTCCACCAAGACGGAATCGTTACCGAGATCTCAAGCTGGAATAATCTCTC 681

DB 605 TGCACTAAGTCCACCAAGACGGAATCGTTACCGAGATCTCAAGCTGGAATAATCTCTC 664
QY 682 TAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAAGA 741
DB 665 TAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAAGA 724
QY 742 GCAAGTTCTTCAGAGCTTCTGTGGAGCCCTCTTACGCTCGCTGAGATGATCAAG 801
DB 725 GCAAGTTCTTCAGAGCTTCTGTGGAGCCCTCTTACGCTCGCTGAGATGATCAAG 784
QY 802 GGAAGCCCTTACAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 861
DB 785 GGAAGCCCTTACAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 844
QY 862 TGTGTCATGAGCACCATGCTTTGACCGGAGAGATCAATAACAATGAGAGCAATCA 921
DB 845 TGTGTCATGAGCACCATGCTTTGATGAGGAGATGATCAATAAGTGTGAACAGATCA 904
QY 922 GTAAAGGAGCTTACCGTGAAGCGGACCAAGCGGCTGATGCTGTGGCTGATCGGTGGC 981
DB 905 GCAAGCGGAGCTTACCGGAGAGCACTAAACCTCTGATGCTGTGGCTGATCGGTGGC 964
QY 982 TGTTAATGATGGAACCCCAAGCTGTGGGACCACTGAGAGATGATGATGATGATGATG 1041
DB 965 TGTTAATGATGGAACCCCAAGCTGTGGGACCAAGCTGAGAGATGATGATGATGATG 1024
QY 1042 TCATCTGAGGATGACCAAGCAGAGTGGGAGACGAGAGGCTGAGAGGAGTGAAGGAG 1101
DB 1025 TCATCTGAGGATGACCAAGCAGAGTGGGAGACGAGAGGCTGAGAGGAGTGAAGGAG 1084
QY 1102 CTAAGTCTGCTGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 1161
DB 1085 CTAAGTCTGCTGCGGACCACTTCACCTGCGGCTCCCGACCATGAGTGGCTTTAC 1144
QY 1162 TCTGAGAGATGAGACCAAGTGTGAGCTTCTTCAAGACAGACAGTGGGAGATGGA 1221
DB 1145 TCTGAGAGATGAGACCAAGTGTGAGCTTCTTCAAGACAGACAGTGGGAGATGGA 1204
QY 1222 GCACTGATCTGAGGCTGAGAGGAGCAATCTCTTAAGAGTCCGGAAGAGATGGA 1281
DB 1205 GCACTGATCTGAGGCTGAGAGGAGCAATCTCTTAAGAGTCCGGAAGAGATGGA 1264
QY 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTGCAAGACA 1341
DB 1265 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTGCAAGACA 1324
QY 1342 GCTTAAGTCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTGCAAGACA 1401
DB 1325 GCTTAAGTCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTGCAAGACA 1384
QY 1402 AGAGAGACCTTCAAGAGCTGAGACCGGCTGATCTCTCAAGAGGACGCTGCTGCTG 1461
DB 1385 AGAGAGACCTTCAAGAGCTGAGACCGGCTGATCTCTCAAGAGGACGCTGCTGCTG 1438
QY 1462 TATCCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTGCAAGAGGCAATCTGTTACT 1521
DB 1439 TATCCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTGCAAGAGGCAATCTGTTACT 1495
QY 1522 ACTCTCTTCCCAAGGAAAGGCAATCTTAAGAGTCTGCAAGAGGCAATCTGTTACT 1581
DB 1496 ACTCTCTTCCCAAGGAAAGGCAATCTTAAGAGTCTGCAAGAGGCAATCTGTTACT 1555
QY 1582 GTGGGAGACCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1641
DB 1566 GTGGGAGACCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1615
QY 1642 GCATTTCTCAAACTCAATGAGCAAGTCTTCCGCAAGAGCTTCAAGAGGCAATCTTACGA 1701
DB 1616 GCATTTCTCAAACTCAATGAGCAAGTCTTCCGCAAGAGCTTCAAGAGGCAATCTTACGA 1675
QY 1702 CTTTGGCTGCTGAGACCAAGTGTGCTTCCATCTGAGAGCGGCGGACGCGGCTT 1761
DB 1676 CTTTGGCTGCTGAGACCAAGTGTGCTTCCATCTGAGAGCGGCGGACGCGGCTT 1735

QY 1762 CAGGGGCTGTGAGTGAAGACGACATCTGTCTCCGAGTCTTTGACCAATTGACCTTGC 1821
| | | | |
DB 1736 CAGGGGCTGTGAGTGAAGACGACATCTGTCTCTGAGTCTTTGACCAAGCTGACCTTGC 1795
| | | | |
QY 1822 CTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACAACCTGAGGGGC 1881
| | | | |
DB 1796 CTGAACGCTCTCCAGAGACCCCACTGCGGGGCTGTGTCTGTGAGCAACAACCTGAGGGGC 1855
| | | | |
QY 1882 TTGAGCAGCTCTCCCTCAGAG-----GTCTGAGACGATGTGTGAGGAGATCTT 1929
| | | | |
DB 1856 TTGAGGAGCCCTCCCTCAGAGGGCCCTGAGACCTGCTGAGGCGCTGGCGGAGATCCTT 1915
| | | | |
QY 1930 TGGGGGATGAGTGTCTTTTCTGTGACAGCTGCCAAGAGTACCTGACCTTACAGACAG 1989
| | | | |
DB 1916 TGGGGGACAGCTGTCTTTCTCTGACAGCTGCCAGAGGATGACAGCACTTACCGACAGG 1975
| | | | |
QY 1990 CCCAGAGATCTGTCTCAAAAGCTCAGCTGAGAGGGAGATGTGCGCCAGTATGAGGGAG 2049
| | | | |
DB 1976 CACTGAGGCTGTCTCAAAAGCTCAGCTGAGTGAAGTGAAGGATTTGCCCGAG-CCCGGTCAG 2034
| | | | |
QY 2050 GCTCTGAGAGGTTTGACAGAGAACCTGGGTGCGATTCTCC 2092
| | | | |
DB 2035 GCTCTCAGATGACAGCTGTGTGACCCCGAGGGAGATGCTTTC 2077
| | | | |
RESULT 13
AK074830 3395 bp mRNA linear PRI 03-SEP-2002
LOCUS AK074830 Homo sapiens cDNA FLJ90349 f18, clone NT2RP2003179, weakly similar
DEFINITION to PUTATIVE SERINE/THREONINE-PROTEIN KINASE EMK (EC 2.7.1.1.-).
ACCESSION AK074830.1 GI:22760531
VERSION AK074830.1
KEYWORDS oligo cloning; f18 (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Iisogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3395)
AUTHORS Iisogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Iisogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo; Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1. .3395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2003179"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RP2"
/note="Cloning vector: pMB18SPJ3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"
118. .2004

/note="unnamed protein product"
/codon_start=1
/protein_id="Bac11234.1"
/db_xref="GI:22760532"
/translation="MESLVPARRSGPTPSAELARPLAELIKSPKLMKQAVRRH
HKINLRHRYFEFLTGKYGKXKARESSGRIVAIKIRKDKIKDEQDLMIHREIE
INSLNPHITAIHVEPENSKIYVNEVYSRDLVDYISEROOLSRERARFRQIV
SAVHYCHONRVHRDLKLENTLIDANGNTKIDPFGSLNLTGQHFDTGCSPLTASB
EIVNGKPYTGPEDVSWSLGLVLIIVHNGVFPQGHDKILVKQISNAYRPPKPSDA
CGILRMILMNPTRRATLEBDVAHMMVNMGYARVGEQEPHSGHSGDSASAD
MLRRSRPLLENKAYKVSFEKQAPGGSTTPELEQHSLSKSKENDMOSHSDDA
DTPAHRGKSNLTPKGIKTKVYASABEVOEPPELSP1PASPGQAPLPPKGIK
KPRQRESGYSSSPSPSESGELIDAGVFPVSGDKEQKPPASGLIHKRIKLNKGF
SQIALBIAPFTTGSIDELAPPRPLPAASPPSASVDSITLSSFPQDLPERLSP
PLRGCVSVNLTGLEBPPSPSGCLRRWQDPLGDSFSLTDCQEVATYRQALRVC
SKLT"

ORIGIN

Query Match 47.4%; Score 1375.8; DB 9; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;

QY 82 CTAATGCTGCCGACCACTTCACTGCGCGTCCCGGACCATGAGTGGACCTTAC 141
| | | | |
DB 77 CTAATGATTCCTCCCTGCGCGCTTGTCTACCTCTGCTGCGCATGAGTCCCTGTTTTCG 136
| | | | |
QY 142 TCAGCGCCGAGCCAGCGCTCCCTGCGCTCCGCTTGGCTCGAGAGCGCCGCGCGC 201
| | | | |
DB 137 CGCGCGCTCCGCGCGCCACTCCCTCGCGCCGAGAGCTA-----GCCGCGCGC 184
| | | | |
QY 202 TGGCGGACGGGCTCATTAATCGCTTAACTCTGATGAAAGACAGCGGTGAAGCGGC 261
| | | | |
DB 185 TGGCGGAGGGCTGATCAAGTCCGCAAGCCCTTAAGAAAGACAGCGGTGAAGCGGC 244
| | | | |
QY 262 ACCATCAAAACAAACCTCGGCAACCGCTACAGAGTCTCTGAGACGCTGGGACAGGGCA 321
| | | | |
DB 245 ACCACCAAGACCAACCTCGGCAACCGCTACAGAGTCTCTGAGACGCTGGGACAGGGCA 304
| | | | |
QY 322 CTAACGGGAAGTGAAGAGGACAGAGAGCTCGGGGCTCTGTGTGCGCATCAAGTCA 381
| | | | |
DB 305 CTAACGGGAAGTGAAGAGGACAGAGAGCTCGGGGCTCTGTGTGCGCATCAAGTCA 364
| | | | |
QY 382 TCAGGAAAGCAAAATTAAGATGAGAGGATCTGTGCAATACGAGGAGATTGAGA 441
| | | | |
DB 365 TCCGGAAGGACAAATTAAGATGAGAGGATCTGTGCAATACGAGGAGATTGAGA 424
| | | | |
QY 442 TCATGCTTCACTCAACCAACCCCAATCATTTGCGATCCATGAAGTTTGAATGCA 501
| | | | |
DB 425 TCATGCTATCATCAACCAACCCCTCAATTCATTCATCAAGAGTTTGAAGACGCA 484
| | | | |
QY 502 GCAAGATTGTGATTTGTCATGAGATGACGAGCGAGCGATCTGTATGATTACATCAGTG 561
| | | | |
DB 485 GCAAGATTGTGATTTGTCATGAGATGACGAGCGAGCGAGCGATCTGTATGATTACATCAGTG 544
| | | | |
QY 562 AGCGGCAACGCTGAGTGAAGCGGAGCGGACGATTTCTTCCAGACAGATGTGTGCC 621
| | | | |
DB 545 AGCGGCAACGCTGAGTGAAGCGGAGCGGAGCGGATTTCTTCCAGACAGATGTGTGCC 604
| | | | |
QY 622 TGCATCTACGCGCCAGGAAGGAGTGTCAACGAGATCTCAAGCTGAGAAACATCTTTC 681
| | | | |
DB 605 TGCATCTATTCATCGAAGAGAGTGTCCACGAGATCTCAAGCTGAGAAACATCTTTC 664
| | | | |
QY 682 TAGATGCAATGGAACATCAAGATTGCTGACTTTGGCCCTTCCAACTTACACAAAG 741
| | | | |
DB 665 TAGATGCAATGGAATATCAAGATTGCTGACTTTGGCCCTTCCAACTTACACAAAG 724
| | | | |
QY 742 GCAAGTTCTTCCAGACGTTTGTGGAGCCCTCTTACGCTTGCCTGAGATAGTCAAG 801
| | | | |
DB 725 GCAAGTTCTTCCAGACATTTGTGGAGCCCTCTTACGCTTGCCTGAGATAGTCAATG 784
| | | | |
QY 802 GGAAGCCCTATGTGGGCGGAGAGTGAACAGCTGTCTGTGGCGCTTCTCTACATCC 861
| | | | |
DB 785 GGAAGCCCTTACACAGCGCCAGAGTGAACAGCTGTCTGTGGCGCTTCTCTACATCC 844
| | | | |

QY 862 TGGTGCATGACACATGATCCCTTTGACGCGAGATCATATAAACAATGATGAAGCAATCA 921
 Db 845 TGGTGCATGACATGATCCCTTTGATGAGGATGATCATTAAGATCTGATGAAGCAATCA 904
 QY 922 GTTAACGGGGCTTACCGTGAAGCCGCCCAAGCGTCCGATGCTGTGGCTGATCCGGTGC 981
 Db 905 GCAACGGGGCTTACCGGAGAGCACTTAACCTCTGATGCTGTGGCTGATCCGGTGC 964
 QY 982 TGTAAATGATGAACCCGACCGGTGCGGCACTGAGAGATGATGACAGATCATTTGGGG 1041
 Db 965 TGTGATGATGAACCCGACCGGTGCGGCACTGAGAGATGATGACAGATCATTTGGGG 1024
 QY 1042 TCAACTGGGGTTACACCAACCGAGTCCGGGAAACAGAAAGCCCTGCTGAGGGTGGGACC 1101
 Db 1025 TCAACTGGGGTTACACCAACCGAGTGGAGAGAGAGAGGCTCCGATGAGGGTGGGACC 1084
 QY 1102 CTAGTGTGATCTTTGGCGGGCTCCATGCGGAGCTGTTAGTGTGCTCTCCGCGCCCC 1161
 Db 1085 CTGGCAGTGAATCTGCGCGGCTCCATGCTGCTGCTCCGCGCTCCGCGCCCC 1144
 QY 1162 TCCGTGAGATGAGCAAGGTGTCAGCTTTCAAGACAGATGCTCCGCGAGGTGGA 1221
 Db 1145 TCCGTGAGATGAGCAAGGTGTCAGCTTTCAAGACAGATGCTCCGCGAGGTGGA 1204
 QY 1222 GCACTGTACCTGGGCTGAGCGGCACTTCTTAAGAAAGTCCGAAAGAGATGACA 1281
 Db 1205 GCAACACCCCTGGCGCTGAGCGGCACTTCTTAAGAAAGTCCGAAAGAGATGACA 1264
 QY 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCTTCTTCTGCGCTTGGCAAGACA 1341
 Db 1265 TGGCGCAGTCTCTTCCACAGTGAACCGGCTGATGACCTGCCATGCGCTTGGCAAGACA 1324
 QY 1342 GCTTTAAGCTTCCGAAAGGATTTCTCAAGAAAGTCTCTCACTGCTGAGGGAGGTAC 1401
 Db 1325 ACTTCAAGCTGCAAGAGGATTTCTCAAGAAAGTCTCTGCAAGAGGAGGTAC 1384
 QY 1402 AGGAGGACCTTCAAGAACTGACCGGCTGCTGATCTCAAGGACGCTGCTGCTG 1461
 Db 1385 AGGAGGACCTTCAAGAACTGACCGGCTGCTGATCTCAAGGACGCTGCTGCTG 1448
 QY 1462 TATCCCTGCTCCAGAAAGGATCTTCAAGAGTCTGACAGCTGAAATCTGCTTACT 1521
 Db 1439 ---CCCTGCTCCCAAGAGGATTTCTCAAGAAAGGCTCCGACAGGAGTCTGGCTACT 1495
 QY 1522 ACTTCTTCCAGAGGACGAGAGTCTGGGAACTTTAGACGCGAGTGTGTTGTGA 1581
 Db 1496 ACTTCTTCCAGAGGACGAGAGTCTGGGAACTTTAGACGCGAGTGTGTTGTGA 1555
 QY 1582 GTGGGAGCCCGTGGAGACAGAGTCTCAAGGCTTCAAGGCTCTCTCCACCGCAAG 1641
 Db 1556 GTGGGAGTCCAGAGACAGAGTCTCCAGAGTCTCAAGGCTCTCTCCATGCAAG 1615
 QY 1642 GCATTTCAAACTCAATGAGCAAGTCTCCGACAGGCTTGAAGGACCTTACCTTGA 1701
 Db 1616 GCATTTCAAACTCAATGAGCAAGTCTCCGACAGGCTTGAAGGACCTTACCTTGA 1675
 QY 1702 CTTTGGCTCTGAGCAACTGAGCTCTCCATCTGACAGCCGCGCCAGCCGCT 1761
 Db 1676 CTTTGGCTCTGAGCAACTGAGCTCTCCATCTGAGAGCCGCGCGCCAGCCGCT 1735
 QY 1762 CAGGGGCTGAGAGAGAGAGAGTCTGTCCTCCAGAGTCTTGAACAATGAGCTTGC 1821
 Db 1736 CAGGGGCTGAGAGAGAGAGAGTCTGTCCTCCAGAGTCTTGAACAATGAGCTTGC 1795
 QY 1822 CTGAACGCTTCTCCGAAACCCCACTGAGGGCTGTGTCTGTGAGCAACAACCTGAGGGGC 1881
 Db 1796 CTGAACGCTTCTCCGAAACCCCACTGAGGGCTGTGTGTGTGAGCAACAACCTGAGGGGC 1855
 QY 1882 TTGAGAGGCTCTCTCAAGAG-----GTGGAAGCATGTGTGAGAGAAATCTT 1929
 Db 1856 TTGAGAGGCTCTCTCAAGAGGCTCTGAGAGGCTGTGAGAGGCTGTGAGAGATCTT 1915

QY 1930 TGGGGATGACCTGCTTTCTCTGACAGATGCTCCAGAGGCTGATGCTGACCTTACAGACAG 1989
 Db 1916 TGGGGATGACCTGCTTTCTCTGACAGATGCTCCAGAGGCTGATGCTGACCTTACAGACAG 1975
 QY 1990 CCTGAAGATCTGCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2049
 Db 1976 CACTGAGAGGCTGCTCAAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034
 QY 2050 GCTCTGAGAGGCTTTCAG 2092
 Db 2035 GCTCTGAGAGGCTTTCAG 2077

RESULT 14
 BC017306
 LOCUS
 DEFINITION
 Homo sapiens likely ortholog of rat SNR1AMP-activated protein kinase, mRNA (CDNA clone MGC:29812 IMAGE:5088037), complete cds.
 BC017306
 VERSION
 ACCESSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3402)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K., Hopkins, R.F., Jordan, H., Moore, J., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheltz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huylk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalke, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marz, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3402)
 Strausberg, R.
 Direct Submission
 Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:16878205.
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Lou Staudt
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc@nhgri.nih.gov
 Ahlter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hashiguchi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maekari, B., Maestrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,

Teurjeon, C., Vogt, D.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 39 Row: 0 Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 13565921.

FEATURES

Source

Location/Qualifiers
1..3402
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:29812 IMAGE:5088037"
/issue_type="lymph, lymphoma"
/clone_id="NH MGC_99"
/lab_host="DH10B"
/note="Vector: pOTB7"
1..3402
/gene="SNARK"
/note="synonym: DKFZP434J037"
/db_xref="LocustID:81788"
/db_xref="MIM:608131"
100..1986
/gene="SNARK"
/codon_start=-1
/product="hypothetical protein DKFZP434J037"
/protein_id="AAH17306.1"
/db_xref="GI:16878206"
/db_xref="LocustID:81788"
/db_xref="MIM:608131"

CDS

100..1986
/gene="SNARK"
/codon_start=-1
/product="hypothetical protein DKFZP434J037"
/protein_id="AAH17306.1"
/db_xref="GI:16878206"
/db_xref="LocustID:81788"
/db_xref="MIM:608131"
/translation="MESLVAFRBSGTPSAELARPIAELGILSPKPLMKQAVKRHH
HKNNRRYFLETLGKTYGKAPKABESSRLVAKISIRDKIDKDDMLHIREIE
IMSLNPHILAIHEVPENSSKIVMEVASRDLVDYISERQDSREARHFRQIV
SAVHYQONVHNDLKENLILVANGENIKIADGSLNLYHQGLQTFGSPLYASB
EIVNGKYTPEDVMSGLVLLYLVAHGTPEFDGDKILVKQISNCAVYRPPKPSA
CGILRLMNTPRATLEDAVSMWVWGYATRVGROBA PHEGHPGSDASASMD
WLRSSPILNKAIVCSFKQAHFGSGTTPGLERHSLSKSKKNDMAQSLSDPA
DTPAHRGKSNLKPKEILKKTVSASABGVQEDPPELSPITASFGQAPLLPKGIILK
KPRQESGYISSPEPSGSLDGLADVFSGDPEQKRPQASGLILHRKQILILKNGF
SQRLBLAATTFGSLDELAPRPLARASPSGASDSILSSSPQLDPERLPEP
PLRGCVVDNLITGLEBPSBSPGSLARMQDPLGDSCFSLDQEVATATYRQALRYC
SKLT"

ORIGIN

Query Match 47.4%; Score 1375.8; DB 9; Length 3402;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;

82 CTACTGTCGCCGACCCCACTGCACTGCGGCTCCCCGACCATGAGTCCGTGCTTAC 141
Db CTACTGATTCCTCCGCGCCCTTGCTCACTCTGCTGCGCATGAGTCCGTGCTTTCG 118
142 TCCAGCCGCCGACGAGCTCCCTCGGCTCCGCGCTGCGCTCGAGAGCGCCGCGCGC 201
Db TCCAGTATTCCTCCGCGCCCTTGCTCACTCTGCTGCGCATGAGTCCGTGCTTTCG 118
119 CGGCGGCTCTCGGCGCCCACTCTCTCGGCGCGAGACTA-----GCCGCGCGC 166
202 TGGCGAGCGGCTCATCAAGTCCTAACTCTGATGAGAGACAGCGGTGAGAGCGC 261
Db TGGCGAGAGGGCTGATCAAGTCGCGCAAGCCCTAATGAGAGACAGCGGTGAGAGCGC 226
262 ACCATCAACAACAACCTGCGGACCGGCTACGAGTTCTGAGAGACGCTGGGCGAGGCA 321
Db ACCACCAACAACAACCTGCGGACCGGCTACGAGTTCTGAGAGACCTTGGGCGAGGCA 286
322 CCTACGGGAAGTGAAGAAGCAAGAGAGCTGGGGCGTCTGAGTCAAGTCA 381
Db CCTACGGGAAGTGAAGAAGCGCGGAGAGCTGGGGCGCTGAGTCAAGTCA 346
382 TCAGAAGAACAACAATCAAGATGAGAGAGATCTGTGACATACGAGGAGATTGAGA 441
Db TCGGAAGAACAACAATCAAGATGAGAGAGATCTGTGACATACGAGGAGATTGAGA 406

442 TCATGCTCTGCAACCAACCCCAATGCGATTCGCAATGCAAGTGTGAGATGCA 501
Db TCATGCTATCACTCAACCAACCTTCAATGCTCAATGCAAGTGTGAGAACAGCA 466
502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
562 AGCGGCAAGGCTGATGAGCGGGAGCGGAGGATTTCTTCCGACAGATGTGTGCCC 621
Db AGCGGCAAGGCTGATGAGCGGGAGCGGAGGATTTCTTCCGAGATGTGTGCCC 586
622 TGCACTATGCGACCAAGAGGAGTGGTTCACGAGATCTCAAGCTGGAAGAACCTCTC 681
Db TGCACTATGCGACCAAGAGGATTTGCTCAAGATCTCAAGCTGGAAGAACCTCTC 646
682 TAGATGCAATGGAACATCAAGATGATGATGATGATGATGATGATGATGATGATG 741
Db TAGATGCAATGGAACATCAAGATGATGATGATGATGATGATGATGATGATGATG 706
742 GCAAGTTCTCCAGACGTTCTGTGGAGCCCTCTTAAGCCTCGCTGAGATGTCAGC 801
Db GCAAGTTCTCCAGACATTTCTGTGGAGCCCTCTTAAGCCTCGCTGAGATGTCAG 766
802 GGAAGCCCTATGAGGCGCCAGAGGTGACAGCTGTGTCTGGGCTTCTCTGATCATCC 861
Db GGAAGCCCTATGAGGCGCCAGAGGTGACAGCTGTGTCTGGGCTTCTCTGATCATCC 826
862 TGGTGATGACACCATGCTCTTGAACGGGAGATCATTAACAACCTGATGAAGCAATCA 921
Db TGGTGATGACACCATGCTCTTGAACGGGAGATCATTAAGATCTGATGAAGCAATCA 886
922 GTAAAGGAGCTTACCGTGAAGCGGCGCAAGCGTTCGATGCTGTGCTGATTCGGTGC 981
Db GTAAAGGAGCTTACCGGAGCGGCGCAAGCGTTCGATGCTGTGCTGATTCGGTGC 946
982 TGTTAAAGTGAACCCCAACCGGTGGGCGCACTGAGAGATGATGACCAATGATGGGG 1041
Db TGTTAAAGTGAACCCCAACCGGTGGGCGCACTGAGAGATGATGACCAATGATGGGG 1006
1042 TCAACTGGGATTCACCAACCGGAGTGGGAGAACAGAGGCTGCGTGAAGGATGGGACCC 1101
Db TCAACTGGGATTCACCAACCGGAGTGGGAGAACAGAGGCTGCGTGAAGGATGGGACCC 1066
1102 CTAGTGTGACTTTGGCGGAGCTTCAATGCGGAGCTGATGCTGCTCGGCGCC 1161
Db CTAGTGTGACTTTGGCGGAGCTTCAATGCGGAGCTGATGCTGCTCGGCGCC 1126
1162 TCTGAGAGATGAGAGCGCAAGTGTGAGCTTTTCAAGAGCAAGCTGCGGAGGTGAA 1221
Db TCTGAGAGATGAGAGCGCAAGTGTGAGCTTTTCAAGAGCAAGCTGCGGAGGTGAA 1186
1222 GCACTGATCTGGGCTGAGAGCGGCAACATTTCTTAAGAGTCCCGAAGAGAGATGACA 1281
Db GCACTGATCTGGGCTGAGAGCGGCAACATTTCTTAAGAGTCCCGAAGAGAGATGACA 1246
1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTTTCTCGGCTTGGCAAGACA 1341
Db TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTTTCTCGGCTTGGCAAGACA 1306
1342 GCTTAAGCTTCCGAAAGGATTTCTCAAGAAAAGTCTTACTGCTCAAGGAGGTAC 1401
Db GCTTAAGCTTCCGAAAGGATTTCTCAAGAAAAGTCTTACTGCTCAAGGAGGTAC 1366
1402 AGGAGAGCCTCAAGAACTGAGACCGGTGCTGATCTCAAGGAGCGCTGCTGCTG 1461
Db AGGAGAGCCTCAAGAACTGAGACCGGTGCTGATCTCAAGGAGCGCTGCTGCTG 1420
1462 TATCCCTGCTCCCAAGAGGAGCATCTTAAAGATCTGACAGCGTGAATCTGTTACT 1521
Db TATCCCTGCTCCCAAGAGGAGCATTTCTCAAGAGGAGCGGAGCATCTGCTGCTACT 1477
1522 ACTCTCTCCAGAGCCAGGAGTCTGGGAACTTTAGAGCGCAAGTATGTGTTGTA 1581

Db 1478 ACTCTCTCCGAGAGGAGTAACTGAGGAGCTCTTGAGAGCAGGCACTGTTTGA 1537
Qy 1582 GTGGGAGACCCGAGAGAGAGTCTCAGAGGCTTCAAGGCTCTCCACCGCAAG 1641
Db 1538 GTGGGAGATCCAGAGAGAGAGGCTCTCGAGGCTTCAAGGCTCTCTCATGCAAG 1597
Qy 1642 GCATTCTCAACTCAATGAGCAAGTTCTCCGACAGAGCTTGAAGGCACTACCCCTAGCA 1701
Db 1598 GCATCTCAAACTCAATGAGCAAGTTCTCCAGACAGGCTTGAAGCTGAGGCTCCCAACA 1657
Qy 1702 CTTTGGCTCTCTGAGCAAGTGGCTCTCTCCATCTGAGAGCCCGGCGCCGCTT 1761
Db 1658 CTTTGGCTCTCTGAGCAAGTGGCTCTCTCCATCTGAGAGCCCGGCGCCGCTT 1717
Qy 1762 CAGGGGCTGAGAGAGAGAGAGAGTCTGAGAGCTCTGAGAGCTTGAAGCAATGAGCTTGC 1821
Db 1718 CAGGGGCTGAGAGAGAGAGAGAGTCTGAGAGCTCTGAGAGCTTGAAGCAATGAGCTTGC 1777
Qy 1822 CTGAACGCTCTTCCGAAACCCCACTGAGAGGCTGAGTCTGAGAGCAACTGAGAGGAGC 1881
Db 1778 CTGAACGCTCTTCCGAAACCCCACTGAGAGGCTGAGTCTGAGAGCAACTGAGAGGAGC 1837
Qy 1882 TTGAGCAGCTCTCTCAGAG-----GTGAGAGCAATGAGTGGAGAGAACTCT 1929
Db 1838 TTGAGAGAGCCCTCTCAGAGAGGCTGAGAGCTGAGAGGCTGAGAGGCTGAGAGCTCTT 1897
Qy 1930 TGGGGAGATGAGTCTTCTCTGAGAGAGTCTGAGAGAGTCTGAGAGCTTCAAGCAAG 1989
Db 1898 TGGGGAGATGAGTCTTCTCTGAGAGAGTCTGAGAGAGTCTGAGAGCTTCAAGCAAG 1957
Qy 1990 CCTGAGAGATCTGCTCAAGAGTCTGAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2049
Db 1958 CACTGAGAGGCTGCTCAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2016
Qy 2050 GCTTGAAGAGGTTTGAAGAGAACCTGAGGCTGAGATCTCTC 2092
Db 2017 GCTTGAAGAGGTTTGAAGAGAACCTGAGGCTGAGATCTCTC 2059

RESULT 15
AX086949 3443 bp DNA linear PAT 09-MAR-2001
LOCUS AX086949 Sequence 901 from Patent WO0112659.
DEFINITION AX086949
ACCESSION AX086949
VERSION AX086949.1 GI:13276064
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. Wiemann, S.
AUTHORS Human dna sequences
JOURNAL Patent: WO 0112659-A 901 22-FEB-2001;
German Human Genome Project (DB)
FEATURES
Source 1..3443
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 47.4%; Score 1374.2; DB 6; Length 3443;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

Qy 82 CTACTGCTGCGGAGCCACCTCAGCTGCGGCTCCCGAGAGAGTGGAGGCTTAC 141
Db 87 CTACTGATTCCTCGCGGCTTGGCTACCTCTGCTGCGAGTGGAGTGGCTTTCG 146
Qy 142 TCCAGCGCCGAGGCTCTCTGCGCTCGGCTTGGCTTGGAGAGCGCCGCGCGC 201

Db 147 CGCGGCGCTCGGCGCCCACTCTCGGCGGAGAGCTA-----GCCGCGCGCC 194
Qy 202 TGGGAGAGGAGCTCATGAGAGTGGCTTAACCTCTGATGAGAGAGAGCGGCTGAGAGCGCGC 261
Db 195 TGGGAGAGGAGCTCATGAGAGTGGCTTAACCTCTGATGAGAGAGAGCGGCTGAGAGCGCGC 254
Qy 262 ACCATCAAAACACACCTTGGGAGCGGCTTACGAGTCTTGGAGAGCGTGGGCAAGGCA 321
Db 255 ACCACCAAAAGACACACCTTGGGAGCGGCTTACGAGTCTTGGAGAGCGTGGGCAAGGCA 314
Qy 322 CTTACGGGAGAGTGAAGAGAGAGAGTCTGGGAGCGGCTTGGGAGAGTGAAGTCAAGTCA 381
Db 315 CTTACGGGAGAGTGAAGAGAGAGAGTCTGGGAGCGGCTTGGGAGAGTGAAGTCAAGTCA 374
Qy 382 TCAGAGAGAGCAAAATCAAGATGAGAGAGTCTGTCACATACGAGAGAGAGTGAAG 441
Db 375 TCGGAGAGAGCAAAATCAAGATGAGAGAGTCTGTCACATACGAGAGAGAGTGAAG 434
Qy 442 TCATGCTTCACTCAACACACCCCACTATGCTCATCTGATGAGAGTGAAGTGAAG 501
Db 435 TCATGCTATCACTCAACACACCCCACTATGCTCATCTGATGAGAGTGAAGTGAAG 494
Qy 502 GCAAGATTTGATTTGATGAGAGTATGAGAGAGAGAGAGTCTGATGATGATGATGATGATGAT 561
Db 495 GCAAGATTTGATTTGATGAGAGTATGAGAGAGAGAGAGTCTGATGATGATGATGATGATGAT 554
Qy 562 AGCGGAGAGAGTGAAG 621
Db 555 AGCGGAGAGAGTGAAG 614
Qy 622 TGCACTATGCTCAACAG 681
Db 615 TGCACTATGCTCAACAG 674
Qy 682 TAGATGCAATGAG 741
Db 675 TAGATGCAATGAG 734
Qy 742 GCAAGTCTCTCAAG 801
Db 735 GCAAGTCTCTCAAG 794
Qy 802 GGAAGCCCTATGAG 861
Db 795 GGAAGCCCTATGAG 854
Qy 862 TGGTGAATGAG 921
Db 855 TGGTGAATGAG 914
Qy 922 GTAAAGGAG 981
Db 915 GCAAGCGGAG 974
Qy 982 TGTAAATGAG 1041
Db 975 TGTAAATGAG 1034
Qy 1042 TCAACTGAG 1101
Db 1035 TCAACTGAG 1094
Qy 1102 CTAGTGTGACTTTGGCGGAGCTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161
Db 1095 CTAGTGTGACTTTGGCGGAGCTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1154
Qy 1162 TCTGAG 1221
Db 1155 TCTGAG 1214
Qy 1222 GCACTGATCTGAG 1281
Db 1215 GCACTGATCTGAG 1274


```
Db      20 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTL-NSSSLQKLFREVRIMKVLNHPN 78
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      79 IYKLFVITETKTLVYMEYASGSEVFDVLVHGRKKEAKRKFQIYSAVOYCHQKFI 138
      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      139 VHRDLKLENILLDADWNIKIADFGFSENEFTFGNKLDLPFGSPPYAAPBLFQGKKYDGPV 198
      180 DWSLGLVLLIYLVHGTMPPDGDHKTLLVQISNGAVREPPKSDAC-GLIRMLMNPTR 238
      199 DWSLGLVLLIYLVHSGSLPFDGQNLKELRERVLRGKRYRIPYMTDCEMLLKPLINPSK 258
      239 RATLEDVASHMWVN 252
      259 RGTLEQIMKDRMWN 272
```

RESULT 11

```
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6452156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 722;
Best Local Similarity 51.2%; Pred. No. 5.5e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

      1 YEFLETLGKGTGKVKKARE-SSGRVIAIKSIKDKIKDQODLHRIREIEMSLNHPN 59
      53 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTL-NSSSLQKLFREVRIMKVLNHPN 111
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      112 IYKLFVITETKTLVYMEYASGSEVFDVLVHGRKKEAKRKFQIYSAVOYCHQKFI 171
      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      172 VHRDLKLENILLDADWNIKIADFGFSENEFTFGNKLDLPFGSPPYAAPBLFQGKKYDGPV 231
      180 DWSLGLVLLIYLVHGTMPPDGDHKTLLVQISNGAVREPPKSDAC-GLIRMLMNPTR 238
      232 DWSLGLVLLIYLVHSGSLPFDGQNLKELRERVLRGKRYRIPYMTDCEMLLKPLINPSK 291
      239 RATLEDVASHMWVN 252
      292 RGTLEQIMKDRMWN 305
```

```
RESULT 12
US-10-274-194-4
; Sequence 4, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```

```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-274-194-4
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 722;
Best Local Similarity 51.2%; Pred. No. 5.5e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

      1 YEFLETLGKGTGKVKKARE-SSGRVIAIKSIKDKIKDQODLHRIREIEMSLNHPN 59
      53 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTL-NSSSLQKLFREVRIMKVLNHPN 111
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      112 IYKLFVITETKTLVYMEYASGSEVFDVLVHGRKKEAKRKFQIYSAVOYCHQKFI 171
      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      172 VHRDLKLENILLDADWNIKIADFGFSENEFTFGNKLDLPFGSPPYAAPBLFQGKKYDGPV 231
      180 DWSLGLVLLIYLVHGTMPPDGDHKTLLVQISNGAVREPPKSDAC-GLIRMLMNPTR 238
      232 DWSLGLVLLIYLVHSGSLPFDGQNLKELRERVLRGKRYRIPYMTDCEMLLKPLINPSK 291
      239 RATLEDVASHMWVN 252
      292 RGTLEQIMKDRMWN 305
```

RESULT 13

```
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 745;
Best Local Similarity 51.2%; Pred. No. 5.8e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

      1 YEFLETLGKGTGKVKKARE-SSGRVIAIKSIKDKIKDQODLHRIREIEMSLNHPN 59
      20 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTL-NSSSLQKLFREVRIMKVLNHPN 78
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
```

Db 79 IVKLEPEVJETEKTLYVMEYASGSEVFDVILVHNGMKKEKARAKAFQIVSAVOYCHOKFI 138
 QY 120 VHRDLKLENIILLDANGNITKIADPGLSNLTHKGFLOTFCGSEPLYASBEIVNGKPYGPEV 179
 Db 139 VHRDLKAENLILLDADNNIKIADFGFSENEFTFGKLTFCGSEPYAA-BELFOGKKYDGPV 198
 QY 180 DSWSLGVLILYLHVGTMPEQGDHKTLYLOISNGAYREPPKPSDAC-GLIRWLLMVPTR 238
 Db 199 DVMSLGVILLIYTVSSGLPPFDGQNLKBLRERVLRGKTRIRPYMTSDCENLLKFLIILNPSK 258
 QY 239 RATEDVASHMWVN 252
 Db 259 RGTLEQIMKDRWMN 272

RESULT 14
 US-09-538-092-1022
 ; Sequence 1022, Application US/09538092
 ; Position 1022

```

? APPLICANT: WENTNER, J Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 6214
? LENGTH: 713
? TYPE: PRT
? ORGANISM: Human
? US-09-949-016-6214

```

RESULT 15
US-09-949-016-6214
; Sequence 6214, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:16:15 ; Search time 197.764 Seconds
(without alignments)
1494.237 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGYGVKVKARE.....VNPTRRATLEDVASHMVMVM 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06 COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07 COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US60 COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	1 PCT-US00-14696-11	Sequence 11, Appl
2	1342	100.0	631	25 US-09-980-464-11	Sequence 11, Appl
3	1342	100.0	631	32 US-10-679-063-17712	Sequence 17712, A
4	1342	100.0	631	32 US-10-679-063-19563	Sequence 19563, A
5	1325	98.7	251	29 US-10-343-514-103	Sequence 103, Appl
6	1313	97.8	630	29 US-10-343-514-103	Sequence 41, Appl
7	1313	97.8	630	32 US-10-679-063-20074	Sequence 20074, A
8	1286	96.6	251	29 US-10-343-514-50	Sequence 50, Appl
9	1283	96.3	628	1 PCT-US02-05095A-1163	Sequence 1163, Ap
10	1293	96.3	628	20 US-09-611-526-2919	Sequence 2919, Ap
11	1293	96.3	628	20 US-09-641-377-902	Sequence 902, Appl
12	1293	96.3	628	25 US-09-963-159-2	Sequence 2, Appl1
13	1293	96.3	628	27 US-10-170-205E-24082	Sequence 24082, A
14	1293	96.3	628	29 US-10-370-715B-640	Sequence 640, Appl
15	1293	96.3	628	30 US-10-423-543-44	Sequence 44, Appl
16	1293	96.3	628	32 US-10-679-063-8699	Sequence 8699, Ap
17	1293	96.3	628	35 US-10-932-349-1502	Sequence 1502, Appl
18	1293	96.3	628	37 US-60-452-680-12460	Sequence 12460, A
19	1293	96.3	628	37 US-60-500-337-1502	Sequence 1502, Ap
20	1293	96.3	672	1 PCT-US03-21730-77	Sequence 77, Appl
21	1293	96.3	672	32 US-10-618-941-77	Sequence 77, Appl
22	1286	95.8	594	29 US-10-311-034-12	Sequence 12, Appl
23	1286	95.8	594	37 US-60-216-605-2	Sequence 2, Appl1
24	1276	95.1	251	29 US-10-343-514-101	Sequence 101, Appl
25	1273	94.9	640	29 US-10-322-281-23	Sequence 23, Appl
26	1268.5	94.5	629	25 US-09-979-167-159	Sequence 159, Appl
27	1230	91.7	616	29 US-10-322-281-26	Sequence 26, Appl
28	1128.5	84.1	661	1 PCT-US03-02588-82	Sequence 82, Appl
29	1128.5	84.1	661	1 PCT-US03-38193-2553	Sequence 2553, Ap
30	1128.5	84.1	661	22 US-09-780-949-2	Sequence 2, Appl1
31	1128.5	84.1	661	22 US-09-780-949-6	Sequence 6, Appl1
32	1128.5	84.1	661	22 US-09-791-531-19112	Sequence 149112, A
33	1128.5	84.1	661	27 US-10-170-205E-24527	Sequence 24527, A
34	1128.5	84.1	661	29 US-10-354-358-82	Sequence 82, Appl
35	1128.5	84.1	661	31 US-10-505-928-690	Sequence 690, Appl
36	1128.5	84.1	661	32 US-10-631-467-583	Sequence 583, Appl
37	1128.5	84.1	661	32 US-10-679-063-26806	Sequence 26806, A
38	1128.5	84.1	661	33 US-10-723-860-2553	Sequence 2553, Ap
39	1128.5	84.1	661	35 US-10-932-349-1564	Sequence 1564, Ap
40	1128.5	84.1	661	35 US-10-932-349-1565	Sequence 1565, Ap
41	1128.5	84.1	661	37 US-60-389-395-2	Sequence 2, Appl1
42	1128.5	84.1	661	37 US-60-500-337-1564	Sequence 1564, Ap
43	1128.5	84.1	661	37 US-60-500-337-1565	Sequence 1565, Ap
44	1120.5	83.5	658	1 PCT-US04-14421-150	Sequence 150, Appl
45	1120.5	83.5	658	27 US-10-144-779-497	Sequence 497, Appl

ALIGNMENTS

RESULT 1
PCT-US00-14696-11
Sequence 11, Application PC/TUS0014696
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Virca, George Duke
APPLICANT: Martin, Unja
TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
FILE REFERENCE: 2923-WO
CURRENT APPLICATION NUMBER: PCT/US00/14696
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 631
TYPE: PRT
ORGANISM: Mus musculus
PCT-US00-14696-11

```
Query Match      100.0%; Score 1342; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
   |||
Db 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
   |||

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
   |||
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
   |||

QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
   |||
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 236
   |||

QY 181 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 240
   |||
Db 237 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 296
   |||

QY 241 TLEDVASHMWVNW 253
   |||
Db 297 TLEDVASHMWVNW 309
   |||

RESULT 2
US-09-980-464-11
; Sequence 11, Application US/09980464
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-US
; CURRENT APPLICATION NUMBER: US/09/980,464
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-980-464-11

Query Match      100.0%; Score 1342; DB 25; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
   |||
Db 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
   |||

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
   |||
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
   |||

QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
   |||
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 236
   |||

QY 181 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 240
   |||
Db 237 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 296
   |||

QY 241 TLEDVASHMWVNW 253
   |||
Db 297 TLEDVASHMWVNW 309
   |||

RESULT 3
```

```
US-10-679-063-17712
; Sequence 17712, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 17712
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-679-063-17712

Query Match      100.0%; Score 1342; DB 32; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
   |||
Db 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
   |||

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
   |||
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
   |||

QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
   |||
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 236
   |||

QY 181 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 240
   |||
Db 237 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 296
   |||

QY 241 TLEDVASHMWVNW 253
   |||
Db 297 TLEDVASHMWVNW 309
   |||

RESULT 4
US-10-679-063-19563
; Sequence 19563, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 19563
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-679-063-19563

Query Match      100.0%; Score 1342; DB 32; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
   |||
Db 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
   |||

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
   |||
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
   |||

QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
   |||
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 236
   |||

QY 181 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 240
   |||
Db 237 SMSLGLVLYILVHGTMPPFGODHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTRR 296
   |||

QY 241 TLEDVASHMWVNW 253
   |||
Db 297 TLEDVASHMWVNW 309
   |||
```

Db 177 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 236
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 240
Db 237 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 236
Qy 241 TLEDVASHMWVNM 253
Db 297 TLEDVASHMWVNM 309

RESULT 5
US-10-343-514-103
; Sequence 103, Application US/10343514
; GENERAL INFORMATION:

APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEBEVRRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 103
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-10-343-514-103

Query Match 98.7%; Score 1325; DB 29; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.1e-109;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 60
Db 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 60
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
Db 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
Qy 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Db 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 240
Db 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 240
Qy 241 TLEDVASHMWVNM 251
Db 241 TLEDVASHMWVNM 251

RESULT 6
US-10-343-514-41

; Sequence 41, Application US/10343514
; GENERAL INFORMATION:
APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEBEVRRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT

CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 41
LENGTH: 630
TYPE: PRT
ORGANISM: RAT
US-10-343-514-41

Query Match 97.8%; Score 1313; DB 29; Length 630;
Best Local Similarity 98.0%; Pred. No. 8.2e-108;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 60
Db 57 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 116
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
Qy 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 236
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 240
Db 237 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRWLMVNPTRRA 296
Qy 241 TLEDVASHMWVNM 253
Db 297 TLEDVASHMWVNM 309

RESULT 7

US-10-679-063-20074
; Sequence 20074, Application US/10679063
; GENERAL INFORMATION:
APPLICANT: Edgeton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52054)B
CURRENT APPLICATION NUMBER: US/10/679,063
CURRENT FILING DATE: 2003-10-02
PRIOR APPLICATION NUMBER: 60/415,758
PRIOR FILING DATE: 2002-10-02
NUMBER OF SEQ ID NOS: 27373
SEQ ID NO 20074
LENGTH: 630
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-679-063-20074

Query Match 97.8%; Score 1313; DB 32; Length 630;
Best Local Similarity 98.0%; Pred. No. 8.2e-108;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 60
Db 57 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLDLHIREIEMSLNHPH 116
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176

QY	121	RRDLKLEINLIDANGIKIADFGLSNLYHKGFLOTFCCSPLIYASPEIVNGKPYVGEVD	180
Db	177	HRDLKLEINLIDASGNIKIADFGLSNLYHKGFLOTFCCSPLIYASPEIVNGKPYVGEVD	236
QY	181	SWSLGVLIYLIIVHGTMFPDGDHKTLYVKOISNGAYREPRPSDACGLIRLWLVNPIRRA	240
Db	237	SWSLGVLIYLIIVHGTMFPDGDHKTLYVKOISNGAYREPCPSDACGLIRLWLVNPIRRA	296
QY	241	TLEDVASHMWNVM	253
Db	297	TLEDVASHMWNVM	309

```

RESULT 8
US-10-343-514-50
? Sequence 50, Application US/10343514
? GENERAL INFORMATION:
? APPLICANT: DRUCKER, Daniel J.
? APPLICANT: ROSEN, Cheryl F.
? APPLICANT: LEFEBVRE, Diana L.
? TITLE OF INVENTION: AMK-R-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
? FILE REFERENCE: DPA-DRUC2/PCT
? CURRENT APPLICATION NUMBER: US/10/343,514
? CURRENT FILING DATE: 2003-01-31
? PRIOR APPLICATION NUMBER: PCT/CA01/01109
? PRIOR FILING DATE: 2001-08-02
? PRIOR APPLICATION NUMBER: US 60/222,650
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: US 60/274,613
? PRIOR FILING DATE: 2001-03-12
? PRIOR APPLICATION NUMBER: CA 2,340,780
? PRIOR FILING DATE: 2001-03-28
? NUMBER OF SEQ ID NOS: 109
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 50
? LENGTH: 251
? TYPE: prt
? ORGANISM: RAT
US-10-343-514-50

```

Query Match	96.6%	Score 1296	DB 29	Length 251
Best Local Similarity	98.0%	Pred. No. 8	2e-107	
Matches	246	Conservative	3	Mismatches 2
				Indels 0
				Gaps 0
QY	1	YEFLETLGKGYGKAYKKAESSGRLVAITSIRKDKIKDEODLIHIREIIMSLSLNPHT	60	
Db	1	YEFLETLGKGYGKAYKKAESSGRLVAITSIRKDKIKDEODLIHIREIIMSLSLNPHT	60	
QY	61	IAIHVEFENS SKIVIVMEYASRGDLVDYISERPRLSERDARHPFEROIVSALHYCHONGIV	120	
Db	61	IAIHVEFENS SKIVIVMEYASRGDLVDYISERPRLENRDARHPFEROIVSALHYCHONGIV	120	
QY	121	HRDLKLENIILLDANGNINIKIADPGLSNLHYHKGFLOTFCCSPLVYASPEIYNGKPYGPEVD	180	
Db	121	HRDLKLENIILLDANGNINIKIADPGLSNLHYHKGFLOTFCCSPLVYASPEIYNGKPYGPEVD	180	
QY	181	SWSLIGVLLIYLIVHGIMPEDGDHKTIVKQISNGAYAEPPKSPDAGCIGIRMLVMNPTBPA	240	
Db	181	SWSLIGVLLIYLIVHGIMPEDGDHKTIVKQISNGAYAEPPKSPDAGCIGIRMLVMNPTBPA	240	
QY	241	TLLEDVYASHMVV 251		
Db	241	TLLEDVYASHMVV 251		

```

RESULT 9
PCT-US02-05095A-1163
; Sequence 1163, Application PC/TUS0205095A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-075 (803)
; CURRENT APPLICATION NUMBER: PCT/US02/05095A

```

```

: CURRENT FILING DATE: 2002-03-05
: PRIOR APPLICATION NUMBER: 09/799,451
: PRIOR FILING DATE: 2001-03-05
: NUMBER OF SEQ ID NOS: 1896
: SEQ ID NO: 1163
: LENGTH: 628
: TYPE: PRT
: ORGNISM: Homo sapiens
: PCT-NO: 050955-1163

```

	Query Match	96.3%;	Score 1293;	DB 1;	length 628;
	Best Local Similarity	95.7%;	Pred. No. 5e-106;		
	Matches 242;	Conservative	6;	Mismatches	5; Indels 0; Gaps 0;
Qy	1	YEFLETTGKGYGVKKAKARSSGRVIAIKSI RKDKIDEDDLTHIRREIETIMSSLNPHHI	:	:	:
Dd	53	YEFLFETLKGKGYGVKKAKARSSGRVIAIKSI RKDKIDEDDLTHIRREIETIMSSLNPHHI	:	:	:
Qy	61	IAIHVEFNSSKIYIVMEYASRGDLYDISERPLSRDAHFPRQIVSALHYCHONGIV	:	:	:
Dd	113	IAIHVEFNSSKIYIVMEYASRGDLYDISERQLSRRAHFPRLQIVSAHYCHGNRVV	:	:	:
Qy	121	HREDLKENILLDANGNKTIADFGLSNLYHGKFLQTFCGSPLVASPEIYNKGPPYVGPEVD	:	:	:
Dd	173	HRDLTKENIILLDANGNKTIADFGLSNLYHQGKFQTFCGSPLVASPEIYNKGPPYTGEVD	:	:	:
Qy	181	SWSLGVLITLYLVHGTMPFDODHTLYVKQISNGAYRPRPKPSDACGLIRMLTMNPTRRA	:	:	:
Dd	233	SWSLGVLITLYLVHGTMPFDGDHKILTVKOISNGAYRPRPKPSDACGLIRMLTMNPTRRA	:	:	:
Qy	241	TLEDVASHMMVNVN 253	:	:	:
Dd	293	TLEDVASHMMVNVN 305	:	:	:

```

RESULT 10
US-09-611-526-2919
Sequence 2919, Application US/09611526
GENERAL INFORMATION:
  APPLICANT: OTA, TOSHIO
  APPLICANT: NISHIKAWA, TETSUO
  APPLICANT: ISOGAI, TAKAO
  APPLICANT: HAYASHI, KOJI
  APPLICANT: ISHII, SHIZUKO
  APPLICANT: KAWAI, YORI
  APPLICANT: WAKAMATSU, AI
  APPLICANT: SUGIYAMA, TOMOYASU
  APPLICANT: NAGAI, KEIICHI
  APPLICANT: KOJIMA, SHINICHI
  APPLICANT: OTSUKI, TETSUKI
  APPLICANT: KOGA, HISASHI
  TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
  TITLE OF INVENTION: AND THEIR USES
  FILE REFERENCE: 08335/0122
  CURRENT APPLICATION NUMBER: US/09/611,526
  CURRENT FILING DATE: 2000-07-07
  PRIOR APPLICATION NUMBER: JP 1999-194486
  PRIOR FILING DATE: 1999-07-08
  PRIOR APPLICATION NUMBER: JP 2000-118774
  PRIOR FILING DATE: 2000-01-11
  PRIOR APPLICATION NUMBER: JP 2000-183765
  PRIOR FILING DATE: 2000-05-02
  NUMBER OF SEQ ID NOS: 4484
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 2919
  LENGTH: 628
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-611-526-2919

```

Query Match	96.3%	Score 1293	DB 20	Length 628
Best Local Similarity	95.7%	Pred. No. 5e-106		
Matches 242	Conservative 6	Mismatches 5	Indels 0	Gaps 0

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 112
Db      61 IAIHEVFNSSKIVIMEVYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVFNSSKIVIMEVYASRGDLVDYISERQOLSREARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 180
      173 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 232
Db      181 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 240
      233 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 292
QY      241 TLEDVASHMWVNV 253
      293 TLEDVASHMWVNV 305
Db
```

RESULT 11

```
US-09-641-377-902
; Sequence 902, Application US/09641377
; GENERAL INFORMATION:
; APPLICANT: WIEMANN, STEFAN
; APPLICANT: GASENHUBER, JOHANN
; APPLICANT: TAMPER, JENS
; TITLE OF INVENTION: HUMAN DNA SEQUENCES
; FILE REFERENCE: 087100/0106
; CURRENT APPLICATION NUMBER: US/09/641,377
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,499
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/156,503
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 1793
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 902
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-641-377-902
```

```
Query Match      96.3%; Score 1293; DB 20; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 112
Db      61 IAIHEVFNSSKIVIMEVYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVFNSSKIVIMEVYASRGDLVDYISERQOLSREARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 180
      173 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 232
Db      181 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 240
      233 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 292
QY      241 TLEDVASHMWVNV 253
      293 TLEDVASHMWVNV 305
Db
```

RESULT 12

```
US-09-963-159-2
; Sequence 2, Application US/09963159
```

```
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roly A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-5001
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-159-2
```

```
Query Match      96.3%; Score 1293; DB 25; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 112
Db      61 IAIHEVFNSSKIVIMEVYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVFNSSKIVIMEVYASRGDLVDYISERQOLSREARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 180
      173 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 232
Db      181 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 240
      233 SMSLGVLLYLVGHGTMPEFGDQDHTLVKQISNGAYRPPKPSDAGCGLIRMLMVNPTERRA 292
QY      241 TLEDVASHMWVNV 253
      293 TLEDVASHMWVNV 305
Db
```

RESULT 13

```
US-10-170-205E-24082
; Sequence 24082, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24082
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-24082
```

```
Query Match      96.3%; Score 1293; DB 27; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDBODLHIREIIMSLSNPHI 112
Db      61 IAIHEVFNSSKIVIMEVYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVFNSSKIVIMEVYASRGDLVDYISERQOLSREARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYTGPEVD 180
```

```
Db 173 HHDKLKENTLIDNNGIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYNGKRYTPEVD 232
Qy 181 SMSLGLVLIYLVHGTMPFDGQDHKTLYKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 240
Db 233 SMSLGLVLIYLVHGTMPFDGHDHKLIVKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305

RESULT 14
US-10-370-715B-640
; Sequence 640, Application US/10370715B
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOFENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseasee
; FILE REFERENCE: F1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 640
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-370-715B-640

Query Match 96.3%; Score 1293; DB 29; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKYKAKRESSGRVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
Db 53 YEFLETGKGTGKYKAKRESSGRVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 112
Qy 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIVSALHYCHONGIV 120
Db 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIVSALHYCHONGIV 172
Qy 121 HHDKLKENTLIDNNGIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYNGKRYTPEVD 180
Db 173 HHDKLKENTLIDNNGIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYNGKRYTPEVD 232
Qy 181 SMSLGLVLIYLVHGTMPFDGQDHKTLYKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 240
Db 233 SMSLGLVLIYLVHGTMPFDGHDHKLIVKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305

RESULT 15
US-10-423-543-44
; Sequence 44, Application US/10423543
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Libermann, Rosena K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
```

```
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Miyoung
; APPLICANT: Williamson, Mark J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP103-023OMNIM
; CURRENT APPLICATION NUMBER: US/10/423,543
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 10/278,036
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 09/711,216
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/205,447
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 10/012,055
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/248,325
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 10/003,690
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/248,893
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 10/217,168
; PRIOR FILING DATE: 2002-08-12
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-423-543-44

Query Match 96.3%; Score 1293; DB 30; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKYKAKRESSGRVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
Db 53 YEFLETGKGTGKYKAKRESSGRVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 112
Qy 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIVSALHYCHONGIV 120
Db 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIVSALHYCHONGIV 172
Qy 121 HHDKLKENTLIDNNGIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYNGKRYTPEVD 180
Db 173 HHDKLKENTLIDNNGIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYNGKRYTPEVD 232
Qy 181 SMSLGLVLIYLVHGTMPFDGQDHKTLYKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 240
Db 233 SMSLGLVLIYLVHGTMPFDGHDHKLIVKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305
```

Search completed: May 11, 2005, 14:37:47
Job time : 199.764 secs